

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 14:43:43 : Search time 2658 Seconds  
(without alignments)  
4532.943 Million cell updates/sec

Title: US-09-680-959-125

Perfect score: 414

Sequence: 1 ctaacgtacagaacagcttg.....gttnggaaagcttgaccagt 414

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Se ad: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	384	92.8	2192	10	MMY15742	Y15742 Mus musculus
2	384	92.8	2329	10	BC030924	BC030924 Mus muscu
3	384	92.8	2428	10	MMBDYSTRO	AJ003007 Mus muscu
4	318.4	76.9	2221	9	BC016655	BC016655 Homo sapi
5	318.4	76.9	2244	9	BC016655	Y15742 Homo sapien
6	318.4	76.9	2247	9	HSDYSTRAO	Y17122 Homo sapien
7	318.4	76.9	2407	9	AF022728	AF022728 Homo sapi
8	318.4	76.9	2573	9	HSDTNB2	Y17118 Homo sapien
9	316.8	76.5	2200	9	AK094494	AK094494 Homo sapi
10	202	48.8	380	10	MMU010206	AJ010206 Mus muscu
11	189.4	45.7	4615	5	FSCYRDYS	L05945 Torpedo cal
12	186.2	45.0	1593	10	AF143542	AF143542 Mus muscu
13	186.2	45.0	1883	10	MMDYSTM32	X55227 M. musculus
14	186.2	45.0	1969	10	AF143544	AF143544 Mus muscu
15	186.2	45.0	2607	10	MMDYSTM24	X59226 M. musculus
16	186.2	45.0	3144	10	AF143543	AF143543 Mus muscu
17	183	44.2	1644	9	HSA9668	AJ009668 Homo sapi
18	183	44.2	1707	9	HSU26742	U26742 Human dysr
19	183	44.2	1710	9	BC005300	BC005300 Homo sapi
20	183	44.2	2419	9	HSU26744	U26744 Human dysr
21	183	44.2	2518	9	HSU46744	U46744 Human dysr
22	183	44.2	2817	9	HSU46745	U46745 Human dysr
23	183	44.2	6066	9	HSM804598	AL833285 Homo sapi
24	178	43.0	204521	2	AC130150	AC130150 Rattus no
25	157.2	38.0	173893	9	AC010150	AC010150 Homo sapi
26	150.8	36.4	2818	3	AF277387	AF277387 Drosophil
27	121.8	29.4	1773	3	CEL131742	AJ131742 Caenorhab
28	100.8	24.3	351	10	MMDBNEX3	279789 M. musculus
29	99.8	24.1	525	9	HUMDTN03	U84531 Human dysr
30	99.8	24.1	15885	9	AC022601	AC022601 Homo sapi
31	99.8	24.1	175885	9	AC068506	AC068506 Homo sapi
32	99.8	24.1	178068	2	AF002411	AF002411 Homo sapi
33	94.2	22.8	62198	2	AC013832	AC013832 Drosophil
34	94.2	22.8	118100	2	AC017911	AC017911 Drosophil
35	94.2	22.8	171705	3	AC007476	AC007476 Drosophil
36	94.2	22.8	18557	3	AC007356	AC007356 Drosophil
37	94.2	22.8	233148	3	AE003822	AE003822 Drosophil
38	88.6	21.4	296	10	MMU010207	AJ010207 Mus muscu
39	85.4	20.6	146004	9	AC019144	AC019144 Homo sapi
40	72	17.4	306	11	DM49G45	270900 D. melanoga
41	60.2	14.5	477	10	MMU010208	AJ010208 Mus muscu
42	44.6	10.8	256	10	MMDBNEX4	279790 M. musculus
43	44	10.6	105838	2	AC106483	AC106483 Rattus no
44	43.2	10.4	43956	3	AF098989	AF098989 Caenorhab
45	43.2	10.4	200509	2	AC006910	AC006910 Caenorhab

# ALIGNMENTS

RESULT 1  
MMY15742  
LOCUS MMY15742 2192 bp mRNA linear ROD 29-MAY-1998  
DEFINITION Mus musculus mRNA for dystrobrevin B, partial.  
ACCESSION Y15742  
VERSION Y15742.1 GI:3127923  
KEYWORDS dystrobrevin B.  
SOURCE Mus musculus.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2192)  
AUTHORS Puca, A.A.  
TITLE Direct Submission  
JOURNAL Submitted (28-NOV-1997) A.A. Puca, T.I.G.E.M, Via Olgettina 58,

```
20132 Milano, ITALY
2 (bases 1 to 2192)
Puccia, A.A., Nigro, V., Piluso, G., Belisio, A., Sampaolo, S.,
Quaderi, N., Rossi, E., Di Iorio, G., Ballabio, A. and Franco, B.
Identification and characterization of a novel member of the
dystrobrevin gene family
FEBS Lett. 425 (1), 7-13 (1998)
98200066
9540997
Location/Qualifiers
1. .2192
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="12"
1. .2192
/feature="MDTN-B"
<1. .>2192
/feature="MDTN-B"
/codon_start=1
/product="dystrobrevin B (MDTN-B)"
/protein_id="CAA75752.1"
/db_xref="GI:3127924"
/db_xref="SPTREMBL:O70585"
/translation="VKLTLTGKNLAAGEPRGSPAPAGKRRARIEEGNKRKTMAE
KRLFTMEQAQFVIRLSTYRTACKLREVQKRNHLVDIWNMLEAPRDNGNLTLDH
STEISVSRLETVISIYQNLKRLSPSTHQISVEQSIISLLNPMVAAYDSEGRGKLTVE
SVKMLATMCGGKMDKRLYIFSQMSDSNGLMFGKLDQFLAEALKPLTAVPEGSPFG
YTEHAVRTCPQKKIMNLMFLDTMADPPQCLVPLMLHRLAHVENVFFHVECSYC
HGESMVFYRQCQCHQNLQCNCFWGHASGASHQNHQKHESSWSPAKKLISHAIS
KSLGCVPSRPHVPEQPEKPLDLHLVPPRLTNMNDTVVSHMSGVPTPTKRLQ
YSQDMPNLADHAIASVYARLQDTRVLDSPLRDEHRLIARYAAPAAEAGNMT
RPTDASFNDANKQORLIAELEKNREILQIOLRLEHEQAQSPPEKQAQNPML
LAELRLRQKDEQMSALQESRRELQVLEGLKLLKQAQTSPTSPHGGGRP
MMPVRSAGSTPTHTGQDSLSGVGGVQEAFAQGTGRNRLNLLVAADSTMTMS
LVKELHSGAEQETKTRGLPPEAHSCPSFCCTGPGSLQAARLTQPPRGAQPM
GSGEGLPEAAAGGPACTSIRIPPSWHGLISAGVQEMEKNRKALKK"
BASE COUNT 556 a 613 c 568 g 455 t
ORIGIN
Query Match 92.8%; Score 384; DB 10; Length 2192;
Best Local Similarity 98.5%; Pred. No. 3.5e-110;
Matches 406; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 3 AAGCTACAGACAGCTTGCAGTTACCGATTGTACAGAGCGATGCAACCTTCATCTTG 62
DB 189 AAGCTACAGACAGCTTGCAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 247
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCGAGACATGCGCTTAACACGCTGGACCACA 122
DB 248 TTGATATCTGGAATATGATCGAAGCTTTCGAGACATGCGCTTAACACGCTGGACCACA 307
QY 123 GCACGGAGATCAGCGGTGTCGCGCTTGAGACCGTCTATCTCGTCCATCTACTATCAGTTG 182
DB 308 GCACGGAGATCA-CGCGTGTCCGCTTGAGACCGTCTATCTCGTCCATCTACTATCAGTTG 366
QY 183 AACAGCGCTTCTTCTACTACAGATCAGCGTGGAGCAGTCCATCAGTCTCTCTACT 242
DB 367 AACAGCGCTTCTTCTACTACAGATCAGCGTGGAGCAGTCCATCAGTCTCTCTACT 426
QY 243 AATTTCATGTCGCGCTTACGACAGTGGCGGAGCGANGTGTGACCGGTGTTTCAGNT 302
DB 427 AATTTCATGTCGCGCTTACGACAGTGGCGGAGCGANGTGTGACCGGTGTTTCAGNT 486
QY 303 AAAGCTATGTTAGCAACATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTTC 362
DB 487 AAAGCTATGTTAGCAACATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTTC 546
QY 363 TCCAGATGTCAGATTCATTCAGGCTTAATGATGTNGGAAAGCTTGACCAAGT 414
DB 547 TCCAGATGTCAGATTCATTCAGGCTTAATGATGTNGGAAAGCTTGACCAAGT 598
RESULT 2
BC030924 2329 bp mRNA linear ROD 07-AUG-2002
Mus musculus, clone MGC:31669 IMAGE:4910166, mRNA, complete cds.
BC030924
BC030924.1 GI:21410283
MGC.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 44 Row: k Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 6681234.
Location/Qualifiers
1. .2329
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:31669 IMAGE:4910166"
/tissue_type="Salivary gland, 10 week old female mouse"
/clone_lib="NCI CGAP_SG2"
/lab_host="DH10B"
/feature="Vector: PCMV-SPORT6"
194. .2002
/codon_start=1
/product="Unknown (protein for MGC:31669)"
/protein_id="AAH30924.1"
/db_xref="GI:21410284"
/translation="MIEGGNKRRTMAEKRLFIEMRAQNEVDVILSTYRTACKLRFV
QKRCNLHLDIWNMLEAPRDNGNLTLDHSTEISVSRLETVISIYQNLKRLSPSTHQI
SYEQSISLLNPMVAAYDSEGRGKLTVEFVKAMLATMCGGKMDKRLYIFSQMSDSNG
LMFGKLDQFLKALKPLTAVPEGSPFGYTEHAVRTCPQKKIMNLMFLDTMADPP
PCLVPLMLHRLAHVENVFFHVECSYCHESMGFRYRQCQCHQNLQCNCFWGHASG
SGAHSNQMKHESSWSPAKKLISHAISKSLGCVPSRPHVPEQPEKPLDLHLV
FPRPLTNMNDTVVSHMSGVPTPTKRLQYSQDMPNLADHAIASVYARLQDTRVL
DPSPLRDEHRLIARYAAPAAEAGNMTRPPTDASFNDANKQORLIAELEKNREI
LQEIQLRLEHEQAQSPPEKQAQNPMLLAELRLRQKDEQMSALQESRRELQV
OLEGLMKLLKQAQTSPTSPHGGGRPMMPVRSAGSTPTHTGHPQDSLSGVGGDVQ
EFAQGTGRNRLNLLVAADSTTMTMSLVKELHSG"
BASE COUNT 588 a 637 c 604 g 500 t
ORIGIN
Query Match 92.8%; Score 384; DB 10; Length 2329;
Best Local Similarity 98.5%; Pred. No. 3.6e-110;
Matches 406; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
QY 3 AAGCTACAGACAGCTTGCAGTTACCGATTGTACAGAGCGATGCAACCTTCATCTTG 62
DB 292 AAGCTACAGACAGCTTGCAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 350
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCGAGACATGCGCTTAACACGCTGGACCACA 122
```

```

|||||
Db 351 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCTTAACACGCTGGACCACA 410
|||||
Qy 123 GCACGGAGATCAGCGGTGTCGCGCTGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182
|||||
Db 411 GCACGGAGATCA-GCGTGTCCGCGCTGAGACCGTCATCTCGTCCATCTACTATCAGTTG 469
|||||
Qy 183 AACAAAGCGCTTCCCTTCTACTACACAGATCAGCGTGGAGCAGTCCATCAGTCTCTACTC 242
|||||
Db 470 AACAAAGCGCTTCCCTTCTACTACACAGATCAGCGTGGAGCAGTCCATCAGTCTCTACTC 529
|||||
Qy 243 AATTTCATGTCGCGCTTCTACTACACAGTAGGAGCAGGAGCAGTCCATCAGTCTCTACTC 302
|||||
Db 530 AATTTCATGTCGCGCTTCTACTACACAGTAGGAGCAGGAGCAGTCCATCAGTCTCTACTC 589
|||||
Qy 303 AAAGCTATGTTAGCAACCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 362
|||||
Db 590 AAAGCTATGTTAGCAACCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 649
|||||
Qy 363 TCCAGATGTCAGATTCCTCAATGGCTTAATGATGTTGGAAGCTTGGACCACT 414
|||||
Db 650 TCCAGATGTCAGATTCCTCAATGGCTTAATGATGTTGGAAGCTTGGACCACT 701

```

```

RESULT 3
MMBDYSTRO 2428 bp mRNA linear ROD 19-MAR-1998
LOCUS Mus musculus mRNA for beta-dystrobrevin.
DEFINITION
ACCESSION AJ003007
VERSION AJ003007.1 GI:2980672
KEYWORDS beta-dystrobrevin.
SOURCE Mus musculus
ORGANISM Mus musculus

```

```

REFERENCE
AUTHORS Blake,D.J., Nawrothski,R., Loh,N.Y., Gorecki,D.C. and Davies,K.E.
TITLE beta-dystrobrevin, a member of the dystrophin-related protein family
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (1), 241-246 (1998)
PUBLINE 98081858
MED 9419360
F. . . Location/Qualifiers
1. .2428
/organism="Mus musculus"
/db_xref="taxon:10090"
127..1974
/codon_start=1
/product="beta-dystrobrevin"
/db_xref="GI:2980672"
/db_xref="SPTREMBL:O70563"
/translation="MIEGCKRKTMAEKRLQETEMPAQNFVDVIRLSTYRTACKLRFP
QKCNLHLVDIWNWIAFRNGNLTLDSHTEISVSRLETVISSIIYOLNRLPSTHOI
SVQSISLLNFWAAYDSEGRKLTAVFVKAMLATCGGKMLDKLRYIFQSMDNSG
LMFGLDKLQFLKLTAVFESGFTYTHAVRTCFPOQKIMLNMFLDTMMADPP
PQCLVLPMLHRLAHVENVFPHVSCYCHCESMMGFYRCQCHNYOLCONCFWRGHA
SGAHSNOMHOMKSHSKSPAKLSHAISKLSGCVPSPEPPHVPPEQPEPLDLHLV
PPRLNMDVTYVSHSSGVTPTKRLQYSQDMPNLLADEHALIASYVARLQCHTRVL
DSFRLDEHRLRLTARLAARAAAGNMRPTDASFNFDAKQORLQIALENKREI
LQIQRLLHQEAQFTPEKQNMPLLAELRLQRKDEQLQMSALQESRELMV
QLEGLMKLKEEQKAAQATGSPHTSPTHGGGRPMPPVRSSTAGSTPTHGPDQSL
GVGGDVQEAFAQCTRRNLRLDLVAADSIITNSSLVKELHSEVTPVS"

```

```

BASE COUNT 584 a 686 c 647 g 511 t
ORIGIN
Query Match 92.8%; Score 384; DB 10; Length 2428;

```

```

Best Local Similarity 98.5%; Pred. No. 3.6e-110;
Matches 406; Conservative 0; Mismatches 4; Indels 2; Gaps 2;
Qy 3 AACGTACAGACAGCTTGCAGAGTTACCAAGTTTCCGAGACAATGGCTTAAACACGCTGGACCACA 122
|||||
Db 225 AACGTACAGACAGCTTGCAGAGTTA-CGATTGTACAGAGCGATGCAACCTTATCTCTG 283
|||||
Qy 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCTTAAACACGCTGGACCACA 122
|||||
Db 284 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCTTAAACACGCTGGACCACA 343
|||||
Qy 123 GCACGGAGATCAGCGGTGTCGCGCTGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182
|||||
Db 344 GCACGGAGATCA-GCGTGTCCGCGCTGAGACCGTCATCTCGTCCATCTACTATCAGTTG 402
|||||
Qy 183 AACAAAGCGCTTCCCTTCTACTACACAGATCAGCGTGGAGCAGTCCATCAGTCTCTACTC 242
|||||
Db 403 AACAAAGCGCTTCCCTTCTACTACACAGATCAGCGTGGAGCAGTCCATCAGTCTCTACTC 462
|||||
Qy 243 AATTTCATGTCGCGCTTCTACTACACAGTAGGAGCAGGAGCAGTCCATCAGTCTCTACTC 302
|||||
Db 463 AATTTCATGTCGCGCTTCTACTACACAGTAGGAGCAGGAGCAGTCCATCAGTCTCTACTC 522
|||||
Qy 303 AAAGCTATGTTAGCAACCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 362
|||||
Db 523 AAAGCTATGTTAGCAACCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 582
|||||
Qy 363 TCCAGATGTCAGATTCCTCAATGGCTTAATGATGTTGGAAGCTTGGACCACT 414
|||||
Db 583 TCCAGATGTCAGATTCCTCAATGGCTTAATGATGTTGGAAGCTTGGACCACT 634

```

```

RESULT 4
BC016655 2221 bp mRNA linear PRI 05-NOV-2001
LOCUS Homo sapiens, Similar to dystrobrevin, beta, clone MGC:17163
DEFINITION
ACCESSION BC016655
VERSION BC016655.1 GI:16741717
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 2221)
JOURNAL Strausberg,R.
Direct Submission
Submitted (31-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalob@bcm.tmc.edu.
Villalob, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia,
A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W.,
Muzny,D.M., Gibbs,R.A.

```

```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 13 Row: h Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 15147342.
Location/Qualifiers
1. .2221

```

```

FEATURES
source
1. .2221

```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:17163 IMAGE:3868207"
/tissue_type="Eye, retinoblastoma"
/clone_lib="NIH_MGC_67"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
157..1839
/codon_start=1
/product="Similar to dystrobrevin, beta"
/protein_id="AAH16655.1"
/db_xref="GI:16741718"
/translation="MIEESGNKRKTKMAEKROLFIEMRAQNFEDVIRLSTYRTACKLRFV
QKRCNLHLVDIWNMIEAFRDNGLNTLDTHTSIVSRLETIVISSIYQNKRLPSTHQI
SVEQISILLNFMIAYDSEGRKLTIVFSVKAMLATMGCKMLDKLRYFSQMSDSNG
LMIFSKEQFLKLEVLPTAVPEGSFYTESVTCFPOQKIKMLNFMFLDTMADPP
PQCLVWLPMLRLAHVENVFPVECSYCRCSMMGFYRCQCOHNYQLCQNCFWRGHA
GGPHSNQOMKSHSSWSPAKKLSHAIKSLGCVPTREPPHVPVPEQPEKPLDLAHIV
PPRLTNMNDTWMSHMSGVPTPTKRLSVDLSRDLDEEHLRLIARYAARLAAAGNVTRP
PTDLSFNFDANKQORLAELENKREILQEIQRURLEHQASQPTPEKAQNPVLLA
ELRLRQRMDEQMSALQESRRLMVLQLEMLKLLKAQATGSPHTSPTGGGRPMP
MPVRSAGSTPTHCPCQDSLSGVGDVQEAFAQAEBEGAEERKMQNGDRG"
157..1839
COUNT 597 a 595 c 593 g 436 t
ORIGIN
Query Match 76.9%; Score 318.4; DB 9; Length 2221;
Best Local Similarity 88.6%; Pred. No. 2.1e-89;
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
QY 3 AAGCTACAGAACAGCTTGAAGTTACCGATTTGTACAGAGCGCATGCAACCTTCATCTTG 62
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 255 AACTTACAGAACAGCTTGAAGTTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 313
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCGAGACATGCGCTTAACACGCTGGACACCA 122
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 314 TTGATATCTGGAACATGATTGAAGCTTTCGAGACATGCGCTTAATACACTGGACCATA 373
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 123 GCACGGAGATCAGGCGTGTNCCGCCCTGGAGACCGTCACTCTCCATCTACTATCAGTTG 182
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 374 CCACGGAGATCA-GTGTGTCCGCCCTGGAACCTGCTATCTCTCCATCTACTATCAGTTG 432
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 AACAGCGCTTCTTCTACTACCATGATGAGCGTGGAGAGCCATCCATCTCTCTACTC 242
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 433 AACAGCGCTTCTTCTACTACCAATAGTGTGGAAACATCTATCAGCTTCTCTCTC 492
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 243 AATTTCATGTGCGCGCTACGACAGTGGAGCGCGAGCGANGTTGACCGTGTTCAGNT 302
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 493 AACTTTATGATTGTCATATGACAGTGGGCGGAGGCAAGTTGACGGTATTTTCAGTT 552
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAATGCTGGACAAATTTGAGATACATTTTC 362
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
L 553 AAAGCTATGTTAGCAACCATGTGTGGTGGAAATGCTGGACAAATTTGAGATATGTTTTC 612
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 363 TCCAGATGTCAGATTCCAATGGCTTAATGATGTTNGGAAAGCTTGACCACT 414
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 613 TCCAGATGTCAGATTCCAATGGCTTAATGATGTTNGGAAAGCTTGACCACT 564
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 5
HSDTNB1 2244 bp mRNA linear PRI 27-MAY-1998
LOCUS HSDTNB1
DEFINITION Homo sapiens mRNA for dystrobrevin B DTN-B1.
ACCESSION Y15722
VERSION Y15722.1 GI:3127912
KEYWORDS DTN-B1; dystrobrevin B.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 2244)
Puca.A.A., Nigro,V., Piluso,G., Belsito,A., Sampaolo,S.,
Quaderi,N., Rossi,E., Di Iorio,G., Ballabio,A. and Franco,B.
Identification and characterization of a novel member of the
```

```
dystrobrevin gene family
FEBS Lett. 425 (1), 7-13 (1998)
98200066
PUBMED 9540997
REFERENCE 2 (bases 1 to 2244)
AUTHORS Puca.A.A.
TITLE Direct Submission
JOURNAL 20132 Milano, ITALY
COMMENT Submitted (25-NOV-1997) A.A. Puca, T.I.G.E.M, Via Olgettina 58,
Related sequence Y12712
Location/Qualifiers
1..2244
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="p23"
/tissue_type="skeletal muscle"
1..1884
/gene="DTN-B1"
1..1884
/gene="DTN-B1"
/codon_start=1
/product="dystrobrevin B DTN-B1"
/protein_id="CAA75737.1"
/db_xref="GI:3127913"
/db_xref="SPTREMBL:O60941"
/translation="MIEESGNKRKTKMAEKROLFIEMRAQNFEDVIRLSTYRTACKLRFV
QKRCNLHLVDIWNMIEAFRDNGLNTLDTHTSIVSRLETIVISSIYQNKRLPSTHQI
SVEQISILLNFMIAYDSEGRKLTIVFSVKAMLATMGCKMLDKLRYFSQMSDSNG
LMIFSKEQFLKLEVLPTAVPEGSFYTESVTCFPOQKIKMLNFMFLDTMADPP
PQCLVWLPMLRLAHVENVFPVECSYCRCSMMGFYRCQCOHNYQLCQNCFWRGHA
GGPHSNQOMKSHSSWSPAKKLSHAIKSLGCVPTREPPHVPVPEQPEKPLDLAHIV
PPRLTNMNDTWMSHMSGVPTPTKRLSVDLSRDLDEEHLRLIARYAARLAAAGNVTRP
PTDLSFNFDANKQORLAELENKREILQEIQRURLEHQASQPTPEKAQNPVLLA
ELRLRQRMDEQMSALQESRRLMVLQLEMLKLLKAQATGSPHTSPTGGGRPMPVRSAGSTPTHCPCQDSLS
GVGDVQEAFAQAGTTRRLNLDLLVAADSIITNTMSSLVLELHSAEGBEERKMQNG
DRG"
BASE COUNT 607 a 604 c 577 g 456 t
ORIGIN
Query Match 76.9%; Score 318.4; DB 9; Length 2244;
Best Local Similarity 88.6%; Pred. No. 2.2e-89;
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;
QY 3 AAGCTACAGAACAGCTTGAAGTTACCGATTTGTACAGAGCGCATGCAACCTTCATCTTG 62
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 99 AACTTACAGAACAGCTTGAAGTTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 157
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCGAGACATGCGCTTAACACGCTGGACACCA 122
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 158 TTGATATCTGGAACATGATTGAAGCTTTCGAGACATGCGCTTAATACACTGGACCATA 217
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 123 GCACGGAGATCAGGCGTGTNCCGCCCTGGAGACCGTCACTCTCTCTACTATCAGTTG 182
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 218 CCACGGAGATCA-GTGTGTCCGCCCTGGAACCTGCTATCTCTCTCTACTATCAGTTG 276
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 183 AACAGCGCTTCTTCTACTACCATGATGAGCGTGGAGAGCCATCCATCTCTCTACTC 242
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 277 AACAGCGCTTCTTCTACTACCAATAGTGTGGAAACATCTATCAGCTTCTCTCTC 336
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 243 AATTTCATGTGCGCGCTACGACAGTGGAGCGCGAGCGANGTTGACCGTGTTCAGNT 302
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 337 AACTTTATGATTGTCATATGACAGTGGGCGGAGGCAAGTTGACGGTATTTTCAGTT 396
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAATGCTGGACAAATTTGAGATACATTTTC 362
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 397 AAAGCTATGTTAGCAACCATGTGTGGTGGAAATGCTGGACAAATTTGAGATATGTTTTC 456
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 363 TCCAGATGTCAGATTCCAATGGCTTAATGATGTTNGGAAAGCTTGACCACT 414
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 457 TCCAGATGTCAGATTCCAATGGCTTAATGATGTTNGGAAAGCTTGACCACT 508
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

RESULT 6
HSDYSTRAO          2247 bp  mRNA  linear  PRI 29-MAY-1998
LOCUS              Homo sapiens mRNA for brain dystrobrevin B.
DEFINITION         Y12712
ACCESSION          Y12712.1 GI:2765226
VERSION            dystrobrevin B.
KEYWORDS            Homo sapiens.
SOURCE             Homo sapiens
ORGANISM            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
REFERENCE           Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
AUTHORS            Puca,A.A.
TITLE              Direct Submission
JOURNAL            Submitted (22-APR-1997) A.A. Puca, Seconda Università di Napoli,
INCE              Istituto di Patologia Generale, Larghetto S. Aniello a Caponapoli 2,
JOURNAL            Napoli 80138, ITALY
AUTHORS            2 (bases 1 to 2247)
TITLE              Puca,A.A., Nigro,V., Piluso,G., Belsito,A., Sampaolo,S.,
JOURNAL            Quaderi,N., Rossi,E., Di Iorio,G., Ballabio,A. and Franco,B.
MEDLINE            Identification and characterization of a novel member of the
PUBMED            dystrobrevin gene family
FEATURES            FEBS Lett. 425 (1), 7-13 (1998)
SOURCE             98200066
                    9540997
                    Location/Qualifiers
                    1..2247
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="2"
                    /map="p23"
                    /tissue_type="brain"
                    /dev_stage="adult"
                    184..1887
                    /codon_start=1
                    /product="dystrobrevin B"
                    /protein_id="CAA73249.1"
                    /db_xref="GI:2765227"
                    /db_xref="SPTREMBL:O43782"
                    /translation="MIESGNKRKTKMAEKROLFIEMRAQNFVIRLSTYRTACKLRVY
                    QKRCNLHLVDIWNMIEAFRDNLGNTLDHTTEISVSRLETVISSIYQNLKRLPSTHQL
                    SVEQISILLNFMAIAYDSGRGLTVFSVKAMLATMCGGKMLDKLRVFSOMSDSNG
                    LMIFSKFOFLKVLKLTAVFEGPSFGYTHSVTCFPQQRKIMLNFELTMDADPP
                    PCLVWLPMLHRLAHVENVPEVCYSRCESMGFRYCCQCHNYQIQCNCFWRGHA
                    GGPNSQNHQMEHSWSPAKKLSHAISKSGCVETPREPPHVPFPEQKPLDLAHIV
                    PPRLTNMDTMVMSHSGVPTTKSLVDSRLEDEHRLIARYAARLAAAGNVTRP
                    PTDLNFNDANKQRLIAELNKNREILQETRLRLEHQASOPTPEKAQONPTLLA
                    ELRLQRKDELRMSALQSRRLAYQLLEMLKLEEKQAAQATGSPHTSPTH
                    GGGRPMPFVSTSGSTPTTCHPCQDLSLGVGDVQEAFAQAEEGAEERKQNGKDR
                    G"
BASE COUNT        591 a 606 c 612 g 438 t
ORIGIN
Query Match       76.9%; Score 318.4; DB 9; Length 2247;
Best Local Similarity 88.6%; Pred. No. 2.2e-89;
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

QY 3 ACGGTACAGACAGCTTGCAGTTACCGATTACGAGAGCGATGCAACCTTCATCTTG 62
Db 282 AACTTACAGACAGCGCTGCAAAATTA-CGATTGTACAAAACGATCAACCTTCATCTTG 340
QY 63 TTGATATCTGGATATGATCGAGCTTCCGAGACAAATGCGCTTACACGCTGGACCACA 122
Db 341 TTGATATCTGGACATGATGATGAAGCTTCCGAGACAAATGCGCTTATACACATGGACCAT 400
QY 123 GCACGGAGATCAGCGGTGTNCCGCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182
Db 401 CCACCGAGATCA-GTGTGTCGCGCCTCGAAACTGTCATCTCTCTCATCTACTATCAGTTG 459
QY 183 ACAAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCTCTACTC 242
Db 460 AACAAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCTCTCTC 519

```

Db 260 AACATTACAGACAGCGCTGCAAAATTA-CGATTTGTACAAAAACGATGCAACCTTCATCTTG 318  
QY 63 TTGATATCTGGATATGATCGAGCTTTCCGAGACAATGGCTTACACGCTGGACCA 122  
Db 319 TTGATATCTGGACATGATGAGCTTCCGAGACAATGGCTTATACACTGGACCA 378  
QY 123 GCACGAGATCAGGCGGTGTCGCCCTGGAGACCGTGCATCTGCTCCATCTACTATCAGTTG 182  
Db 379 CCACCGAGATCA-GTGTGTCGCCCTCGAAACTGTCTCTCCATCTACTATCAGTTG 437  
QY 183 AACACGCGCTTCTCTACTACAGATCAGCGTGGAGCAGTCCATCAGTCTCTACTC 242  
Db 438 AACACGCGCTTCTCTACTACCAAAATAGTGTGGAACAATCTATCAGCGCTTCTCTC 497  
QY 243 AATTTTCATGTCGCCCTTACGACAGTGGAGCGGAGGCGAGGCGTGTGACCGTGTTCAGNT 302  
Db 498 AACTTTATGTTGCTGCATATGACAGTGGAGCGGAGGCGAGGCGTGTTCAGTT 557  
QY 303 AAAGCTATGTTAGCAACCATGCTGTGTGGAAAAATGCTGGACAATTTGAGATACATTTTC 362  
Db 558 AAAGCTATGTTAGCAACCATGCTGTGTGGAAAAATGCTGGACAATTTGAGATATGTTTC 617  
363 TCCAGATGTCAGATCCCAATGCTTAATGATTTAGCAAGTTGACCACT 414  
Db 618 TCCAGATGTCAGATCCCAATGCTTAATGATTTAGCAAGTTGACCACT 669  
RESULT 8  
HSDTNB2  
LOCUS Homo sapiens mRNA for dystrobrevin B DTN-B2. linear PRI 27-MAY-1998  
DEFINITION Y15718  
ACCESSION Y15718  
VERSION Y15718.1 GI:3133086  
KEYWORDS DTN-B2; dystrobrevin B.  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 2573)  
AUTHORS Puca,A.A., Nigro,V., Piluso,G., Belisito,A., Sampaolo,S.,  
Quaderi,N., Rossi,E., Di Iorio,G., Ballabio,A. and Franco,B.  
TITLE Identification and characterization of a novel member of the  
dystrobrevin gene family  
JOURNAL FEBS Lett. 425 (1), 7-13 (1998)  
MEDLINE 9820066  
PUBMED 9540997  
REFERENCE 2 (bases 1 to 2573)  
AUTHORS Puca,A.A.  
TITLE Direct Submission  
JOURNAL Submitted (25-NOV-1997) A.A. Puca, T.I.G.E.M. Via Olgettina 58,  
20132 Milano, ITALY  
COMMENT Related sequence Y12712.  
FEATURES  
source  
1..2573  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="p23"  
/tissue\_type="skeletal muscle"  
/clone\_lib="Clontech"  
1..1677  
/gene="DTN-B2"  
1..1677  
/gene="DTN-B2"  
/codon\_start=1  
/product="dystrobrevin B DTN-B2"  
/protein\_id="CAAY5733.1"  
/db\_xref="GI:3133087"  
/db\_xref="SPTREMBL:O60881"  
/translation="MIEESGNKRKTMAEKRLQFLFIEMRAQNFVIRLSTYRTACKLRVF  
QKCNHLVDIWNMIAFRONGLTIDHTTEISVRLTEYVSIYVQLNKLPLSTHOI  
SVQGSILLNLFNIAVDSEGRKLTWFSVKAMLATMCGKMLDKLRYVFSOWSDSNG  
LMIFSFKDFELKVLPLTAVFEGSPFGSYTEHSVRCFPQQRKIMLNMFLDTMMADPP

POCLVWLPMLHRLAHVENPVPVCSYCRCSMMGFYRCOOCHNYQLCONCFWRGHA  
GGPSNQHQKHEHSSWSPKSLSHAISSKSGVPTREPPHVPPEOPEKPLDLAHIV  
PPRLTNNMDTWVSHMSSGVPTPKRQYSDIDPSHLADEHALJASVYARLQHCARVL  
DSPRLDEHRLIARYAARAAEAGNVTPRTDLSFNDANKQROLIAELEENREI  
LQETRLRLHEQASQPTPEKAQONVTLIAELRLROROKDELOMAGLSQSRRLAV  
QLELMKLLBEEKQAVSELNRPPOPSFARSFPHINYSPLPAHHL"  
BASE COUNT 709 a 617 c 565 g 682 t  
ORIGIN  
Query Match 76.9%; Score 318.4; DB 9; Length 2573;  
Best Local Similarity 88.6%; Pred. NO. 2.2e-89;  
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;  
QY 3 AACGTACAGAACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62  
Db 99 AACTTACAGAACAGCTTGCAAAATTA-CGATTTGTACAAAAACGATGCAACCTTCATCTTG 157  
QY 63 TTGATATCTGGATATGATCGAGCTTTCCGAGACAATGGCTTACACGCTGGACCA 122  
Db 158 TTGATATCTGGACATGATGAGCTTCCGAGACAATGGCTTATACACTGGACCA 217  
QY 123 GCACGAGATCAGGCGGTGTCGCCCTGGAGACCGTGCATCTGCTCCATCTACTATCAGTTG 182  
Db 218 CCACCGAGATCA-GTGTGTCGCCCTCGAAACTGTCTCTCCATCTACTATCAGTTG 276  
QY 183 AACACGCGCTTCTCTACTACAGATCAGCGTGGAGCAGTCCATCAGTCTCTACTC 242  
Db 277 AACACGCGCTTCTCTACTACCAAAATAGTGTGGAACAATCTATCAGCGCTTCTCTC 336  
QY 243 AATTTTCATGTCGCCCTTACGACAGTGGAGCGGAGGCGAGGCGTGTTCAGNT 302  
Db 337 AACTTTATGTTGCTGCATATGACAGTGGAGCGGAGGCGAGGCGTGTTCAGTT 396  
QY 303 AAAGCTATGTTAGCAACCATGCTGTGTGGAAAAATGCTGGACAATTTGAGATACATTTTC 362  
Db 397 AAAGCTATGTTAGCAACCATGCTGTGTGGAAAAATGCTGGACAATTTGAGATGTTTC 456  
QY 363 TCCAGATGTCAGATCCCAATGCTTAATGATTTAGCAAGTTGACCACT 414  
Db 457 TCCAGATGTCAGATCCCAATGCTTAATGATTTAGCAAGTTGACCACT 508  
RESULT 9  
AK094494 2200 bp mRNA linear PRI 15-JUL-2002  
LOCUS Homo sapiens cDNA FLJ37175 fis, clone BRACE2028410, highly similar  
DEFINITION to Homo sapiens beta-dystrobrevin (BDTN) mRNA.  
ACCESSION AK094494  
VERSION AK094494.1 GI:21753568  
KEYWORDS Oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens cerebellum cDNA to mRNA, clone\_lib:BRACE2  
clone:BRACE2028410.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,  
Otsuki,T., Sato,H., Wakamatsu,A., Ishii,S., Yamamoto,J., Isono,Y.,  
Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,  
Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,  
Wagatsuma,M., Murakawa,K., Kanehori,K., Takahashi-Fujii,A.,  
Oshima,A., Suzuki,Y., Sugano,S., Nagahara,K., Masubo,Y., Nagai,K.  
and Isogai,T.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2200)  
AUTHORS Isogai,T. and Yamamoto,J.  
TITLE Direct Submission  
JOURNAL Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Kazusa-Kamatari, Kisarazu, Chiba 252-0812, Japan  
COMMENT (E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:





```
QY 4 ACGTACAGACAGCTTGCAGTACCAGTTTGTACAGAGCGATGCAACCTTCATCTTGT 63
DB 235 ACCTACAGGACAGCATGTAGCT-TCGATTTGTGAGAAAAATGCAATGTGCATTTGT 293
QY 64 TGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGCCCTTAACACGCTGGACACAG 123
DB 294 AGATGTTGGAATGTATAGAGCATTCAGAGAAATGGTTAATACATGATCCGAA 353
QY 124 CACGGAGATCAGCGGTGTCGCCCTGGAGACCGTCACTCGTCCATCTACTATCATAGTTGA 183
DB 354 CGCTGAACCTTA-GTGTGGCTGCTGGAAGCTATTATCTCCGCAATATCTTTACCAGCTGA 412
QY 184 ACAAGGCGCTCTCTACTACACAGATCAGCTGGAGAGTCCATCAGTCTCCTACTCA 243
DB 413 ACAACAGAAATGCTTAACCCACCAAGATAACGTTGAGACGTCGGTTAGTTACTGCTGA 472
QY 244 ATTTTCATGTCGCCCTACGACAGTGGAGCGGAGGCGGANGTTGACCGTGTTCAGNTA 303
DB 473 ACTTTCGCTGCTGCTTATGATGAGAGGCGCATGGCAAAATATCAGTTTTTGTGTGA 532
QY 304 AAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTCT 363
DB 533 AAGTGTGTGTAGCAACATATGTCGAGGAAAGATTCTGGATAAATGAGATATATTTCT 592
QY 364 CCCAGATGTCAGATCCCAATGCTTAATGATGCTTNGGAAAGCTTGACCACT 414
DB 593 CCCAAATATCAGACTCCCAATGGGTCATGATAAATGCAAGTTTGTATCAGT 643

RESULT 12
AF143542 1593 bp mRNA linear ROD 19-OCT-1999
LOCUS M.musculus alpha-dystrobrevin 3 mRNA, complete cds.
DEFINITION M.musculus alpha-dystrobrevin 3 mRNA, complete cds.
ACCESSION AF143542
VERSION AF143542.1 GI:4929244
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Enigk,R.E. and Maimone,M.M.
TITLE Differential expression and developmental regulation of a novel
alpha-dystrobrevin isoform in muscle
JOURNAL Gene 238 (2), 479-488 (1999)
MEDLINE 20035753
PUBMED 10570976
REFERENCE 2 (bases 1 to 1593)
AUTHORS Maimone,M.M. and Enigk,R.E.
TITLE Direct Submission
JRNAL Submitted (16-APR-1999) Anatomy and Cell Biology, SUNY Health
Science Center at Syracuse, 750 East Adams St., Syracuse, NY 13210,
USA
FEATURES
source Location/Qualifiers
1..1593
/organism="Mus musculus"
/strain="C3H"
/db_xref="taxon:10090"
/map="between markers D18Mit62 and D18Mit23"
/clone="11.2"
/cell_line="BC3H1"
/tissue_type="muscle"
1..244
245..1360
/codon_start=1
/product="alpha-dystrobrevin 3"
/protein_id="AAD33913.1"
/db_xref="GI:4929245"
/translation="MIEDSGRRGNMFAEROLFAMERQADLDRLSTYRTACKLRFV
QKCNHVDIWNVIALRNNLDNIENLVALEAVLSTFYQLNKRMPPTHQI
HVEQISLLFLAADFEGHGKISFVAKMALATCGGKIMDKLRYIFSMISDSG
VMVGRYDQFLREVLKLPATVAFEGSPFQTSARSCFSQOKKVTNLGFLDLSMDPP"

QY 4 ACGTACAGACAGCTTGCAGTACCAGTTTGTACAGAGCGATGCAACCTTCATCTTGT 63
DB 235 ACCTACAGGACAGCATGTAGCT-TCGATTTGTGAGAAAAATGCAATGTGCATTTGT 293
QY 64 TGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGCCCTTAACACGCTGGACACAG 123
DB 294 AGATGTTGGAATGTATAGAGCATTCAGAGAAATGGTTAATACATGATCCGAA 353
QY 124 CACGGAGATCAGCGGTGTCGCCCTGGAGACCGTCACTCGTCCATCTACTATCATAGTTGA 183
DB 354 CGCTGAACCTTA-GTGTGGCTGCTGGAAGCTATTATCTCCGCAATATCTTTACCAGCTGA 412
QY 184 ACAAGGCGCTCTCTACTACACAGATCAGCTGGAGAGTCCATCAGTCTCCTACTCA 243
DB 413 ACAACAGAAATGCTTAACCCACCAAGATAACGTTGAGACGTCGGTTAGTTACTGCTGA 472
QY 244 ATTTTCATGTCGCCCTACGACAGTGGAGCGGAGGCGGANGTTGACCGTGTTCAGNTA 303
DB 473 ACTTTCGCTGCTGCTTATGATGAGAGGCGCATGGCAAAATATCAGTTTTTGTGTGA 532
QY 304 AAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATGAGATACATTTCT 363
DB 533 AAGTGTGTGTAGCAACATATGTCGAGGAAAGATTCTGGATAAATGAGATATATTTCT 592
QY 364 CCCAGATGTCAGATCCCAATGCTTAATGATGCTTNGGAAAGCTTGACCACT 414
DB 593 CCCAAATATCAGACTCCCAATGGGTCATGATAAATGCAAGTTTGTATCAGT 643

RESULT 13
MMDYSTM32 1883 bp mRNA linear ROD 24-APR-1996
LOCUS M.musculus mRNA for dystrobrevin (clone m32).
DEFINITION X95227
ACCESSION X95227.1 GI:1246784
VERSION 87 kDa protein; dystrobrevin.
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Blake,D.J., Nawrotzki,R., Peters,M.F., Froehner,S.C. and
Davies,K.E.
TITLE Isoform diversity of dystrobrevin, the murine 87-kDa postsynaptic
protein
JOURNAL J. Biol. Chem. 271 (13), 7802-7810 (1996)
MEDLINE 96205975
PUBMED 8631824
REFERENCE 2 (bases 1 to 1883)
AUTHORS Blake,D.J.
TITLE Direct Submission
JOURNAL Submitted (18-JAN-1996) D.J. Blake, Genetics Laboratory, Dep. of
Biochemistry, Univ. of Oxford, South Park Road, Oxford, Oxfordshire
OX1 3QU, UK
FEATURES
source Location/Qualifiers
1..1883
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="m32"
/tissue_type="brain"
/dev_stage="adult"
```

[illegible]

KEYWORDS 87 kDa protein; dystrobrevin.  
SOURCE Mus musculus  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE 1 (bases 1 to 2607)  
AUTHORS Blake,D.J., Nawrotzki,R., Peters,M.F., Froehner,S.C. and Davies,K.E.  
TITLE Isoform diversity of dystrobrevin, the murine 87-kDa postsynaptic protein  
JOURNAL J. Biol. Chem. 271 (13), 7802-7810 (1996)  
MEDLINE 96205975  
PubMed 8631824  
REFERENCE 2 (bases 1 to 2607)  
AUTHORS Blake,D.J.  
TITLE Direct Submission  
JOURNAL Submitted (18-JAN-1996) D.J. Blake, Genetics Laboratory, Dep. of Biochemistry, Univ. of Oxford, South Park Road, Oxford, Oxfordshire OX1 3QU, UK

FEATURES Location/Qualifiers  
1..2607  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="m24"  
/tissue\_type="brain"  
/dev\_stage="adult"  
46..2112  
/codon\_start=1  
/product="dystrobrevin"  
/protein\_id="CAA64518.1"  
/db\_xref="GI:1246783"  
/db\_xref="MGD:MGI:106039"  
/translation="MIEDSGKRGNTMAERQLFAEMRAQDLRIILSTYRTACKLREV  
QKCNLHLVDINVLNRLNLDNIENLVARLEAVLSTIFYOLNKRMPPTHQI  
HVEQISLLNLFLLAEPDPHGKISVFAVKMALATLCGGKINDKLRVIFSMISDSG  
VMYGRVDQLREVLKPLTAVFEGPSGYTEQSNRSCFSQKKVTLNGLDITLMSDPP  
PQCLWLPFLHLRLANVNFHVECSYCHSEMMGFYRCQCHNYQCCDFWRGHA  
GGSHSNQHKQTSWSPAKKLTNLSKSCASSREPLHMPFDPQEPKPLNLAHV  
PPRPVTSMDTLFSHPVSSGSPFITRSMLESSNRLDEHRLIARYAARLAESSSQ  
PTQPSADPISFTIDANKQRLIAELEKNRREILQIRLVEHEQASQTPKAAQO  
NPTLLARLLRQRMSALQSRRLMVQLGLMKLLKEELKOGTRASSPR  
SSPSHTLSRPIPMPIRSASACPTPTHTPDQSLTGVGGVQEAFAQSRRNLRLSLLVA  
ADSLTNTMSLSVLKELNVEASETSEFSRPFEDLAPSPPTSEKAFIAQIHSRAP  
GYIHGGAASTHGDMPDGDPTPTQPDGNYENESYRQLENEQLQLELYLKQLODEAY  
QVSLQG"

CDS

BASE COUNT 723 a 686 c 641 g 557 t  
ORIGIN

every Match 45.08; Score 186.2; DB 10; Length 2607;  
st Local Similarity 68.6%; Pred. No. 1.7e-47;  
Matches 282; Conservative 0; Mismatches 127; Indels 2; Gaps 2;  
QY 4 ACGTACAGACAGCTTGAAGTTCAGGATTCGTACAGAGCGATGCAACCTTCATCTGT 63  
DB 145 ACCTACAGACAGATGCAAGCT-GAGATTGTGCGAAGAAATGCAATTTGCACCTGT 203  
QY 64 TGATATCTGGATATGATCGAAGCTTTCCGAGACAAATGGCCCTTAACACGCTGGACACAG 123  
DB 204 GGACATTTGGAAGCTATTGAAGCATTCGCGAANAACGCTTTGATACCTGGACCCAA 263  
QY 124 CAGGAGATCAGCGGTGTCGCCCTGGAGACCGTCATCTGTCATCTACTATCAGTTGA 183  
DB 264 CATAGAAGTCA-ACGTGCGCGCTGGAGGGGGTCTCCACTATTTTACCAGCTCA 322  
QY 184 ACAAGCGCTTCTTCTACTCACCAGATCAGGTGAGCGAGTCCATCAGTCTCCTACTCA 243  
DB 323 ACAGAGATGCCAACCACTCACCAGATCCAGGTGAGCGAGTCCATCAGTCTCCTGCTGA 382  
QY 244 ATTTTCATGTCGCCGCTACGACAGTGGAGCGCGAGCGAGTGTGACCGTGTGTTTCAGNTA 303  
DB 383 ACTTCTGCTCGCAGCCCTTGACCCGGAAGCCATGGAAATCTCAGTATTGCTGTC 442

Search completed: December 1, 2002, 17:08:47  
Job time : 2665 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 14:38:13 ; Search time 256 Seconds  
(without alignments)  
3641.904.Million cell updates/sec

Title: US-09-680-959-125

Perfect score: 414

Sequence: 1 ctaacgtacagacagcttg.....gttggaagcttgaccagt 414

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Sehed: 2185239 seqs, 1125999159 residues

TOE number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_101002:\*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	318.4	76.9	2247	ABK70286	Human lung cancer
2	183	44.2	2518	AAS73627	DNA encoding novel
3	181.4	43.8	2520	AAS73629	Human lung cancer
4	181	43.7	512	ABK70157	Drosophila melanog
5	150.8	36.4	3045	ABLI4199	Drosophila melanog
6	94.2	22.8	15987	ABLI4198	Human brain Exres
7	81.4	19.7	472	AAQ60778	Human ovarian can
8	78.4	18.9	689	ABL87347	Human secreted pro
9	71.2	17.2	458	AAC01269	

10	39.2	9.5	5106	22	AAS31142	Human diagnostic a
11	38.2	9.2	13815	19	AAV18885	Mus musculus dystr
12	38.2	9.2	13815	24	ABK81960	cDNA encoding mous
13	38.2	9.2	13815	24	ABK199799	Mouse ischaemic co
14	38.2	9.2	19307	17	AAT27558	Shuttle vector pad
15	37.2	9.0	887	24	ABK81993	Human dystrophin C
16	37.2	9.0	1434	24	AAD37243	Human dystrophin r
17	37.2	9.0	1821	24	AAD37241	Human dystrophin r
18	37.2	9.0	2169	24	AAD37232	Human dystrophin r
19	37.2	9.0	3163	21	AAZ48571	A rod shortened dy
20	37.2	9.0	3172	21	AAZ48570	A rod shortened dy
21	37.2	9.0	3446	24	AAD37242	Human dystrophin m
22	37.2	9.0	3510	24	AAD37240	Human dystrophin m
23	37.2	9.0	3531	24	AAD37238	Human dystrophin m
24	37.2	9.0	3747	21	AAZ48566	A rod shortened dy
25	37.2	9.0	3858	24	AAD37237	Human dystrophin m
26	37.2	9.0	3999	24	AAD37234	Human dystrophin m
27	37.2	9.0	4075	21	AAZ48569	A rod shortened dy
28	37.2	9.0	4182	24	AAD37230	Human dystrophin m
29	37.2	9.0	4402	21	AAZ48567	A rod shortened dy
30	37.2	9.0	4402	21	AAZ48568	A rod shortened dy
31	37.2	9.0	4414	24	AAD37260	Adeno-associated v
32	37.2	9.0	4476	24	AAD37259	Adeno-associated v
33	37.2	9.0	4498	24	AAD37258	Adeno-associated v
34	37.2	9.0	4825	24	AAD37257	Adeno-associated v
35	37.2	9.0	4848	24	AAD37263	Adeno-associated v
36	37.2	9.0	4966	24	AAD37256	Adeno-associated v
37	37.2	9.0	4990	24	AAD37262	Adeno-associated v
38	37.2	9.0	5060	24	AAD37264	Adeno-associated v
39	37.2	9.0	5149	24	AAD37255	Adeno-associated v
40	37.2	9.0	5339	24	ABK81998	DNA encoding mini-
41	37.2	9.0	5417	24	ABK81997	DNA encoding mini-
42	37.2	9.0	5462	24	ABK81999	DNA encoding mini-
43	37.2	9.0	5952	22	AAD06794	Human dystrophin g
44	37.2	9.0	8689	24	ABK82000	DNA encoding mini-
45	37.2	9.0	11058	24	AAD37229	Human dystrophin p

#### ALIGNMENTS

##### RESULT 1

ABK70286  
ID ABK70286 standard; cDNA; 2247 BP.  
XX  
AC ABK70286;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE Human lung cancer associated full length cDNA DMSM-52.  
XX  
KW Human; ss; gene; lung cancer; cytostatic; tumour; vaccine.  
XX  
OS Homo sapiens.  
XX  
PN WO200224057-A2.  
XX  
PD 28-MAR-2002.  
XX  
PF 20-SEP-2001; 2001WO-US42232.  
XX  
PR 22-SEP-2000; 2000US-234837P.  
PR 10-OCT-2000; 2000US-239440P.  
PR 29-JUN-2001; 2001US-301928P.  
XX  
(CORI-) CORIXA CORP.  
XX  
Benson DR, Mohamath R, Lodes MJ;  
XX  
WPI; 2002-372001/40.  
XX  
New tumour lung proteins and nucleic acids encoding the proteins, useful  
as vaccines and for treating, preventing, diagnosing or monitoring lung

PT cancer  
PS Claim 1; Page 160-161; 189pp; English.  
XX  
XX The invention relates to an isolated polynucleotide comprising a sequence  
CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),  
CC or their fragments, homologues, variants or complements and their encoded  
CC polypeptides. Also included are an expression vector comprising the  
CC polynucleotide operably linked to an expression control sequence; a host  
CC cell transformed or transfected with an expression vector of; an isolated  
CC polypeptide; a method for detecting the presence of a cancer in a  
CC patient; a fusion protein comprising at least the polypeptide; an  
CC oligonucleotide that hybridises to the polynucleotide under moderately  
CC stringent conditions; a method for stimulating and/or expanding T cells  
CC specific for a tumour protein; an isolated T cell population comprising T  
CC cells prepared from the method of above; a composition comprising a first  
CC component consisting of carriers and immunostimulants, and a second  
CC component selected from the polynucleotides, proteins, antibodies, fusion  
CC proteins, T cell populations and antigen presenting cells expressing the  
CC polypeptide; methods for stimulating an immune response or treating  
CC cancer in a patient by administering the composition and diagnostic kits  
CC comprising at least one of the oligonucleotide of, or an antibody and a  
CC detection reagent consisting of a reporter group. The polypeptides and  
CC polynucleotides are useful as vaccines for the treatment or prevention of  
CC lung cancer, and for diagnosis and monitoring of such cancer. The  
CC polynucleotide, polypeptide and antigen presenting cells can be  
CC used to stimulate or expand T cells specific for a tumorous protein.  
CC The polynucleotides may be used as probes or primers for nucleic acid  
CC hybridisation, and in the preparation of ribozyme molecules for  
CC inhibiting expression of tumour polypeptides and proteins in tumour  
CC cells. The present sequence is one of the 183 lung cancer associated  
CC polynucleotides.  
XX

SQ Sequence 2247 BP; 591 A; 606 C; 612 G; 438 T; 0 other;

Query Match 76.9%; Score 318.4; DB 24; Length 2247;  
Best Local Similarity 88.6%; Pred. No. 2.2e-99;  
Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;  
QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62  
DB 282 AACTTACAGACAGCTGCAAAATTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 340  
QY 63 TTGATATCTGGATATGATCGAGCTTCCGAGACAAATGGCCTTAACACGTGGACCA 122  
DB 341 TTGATATCTGGACATGATGAGCGCTCCGAGACAAATGGCCTTAACACGTGGACCA 400  
QY 123 GCACGGAGATCAGGGCTGTNCCGCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182  
401 CCACCGAGATCA-GTGTGTCGCGCTCGAAGCTGTCATCTCCTCATCTACTATCAGTTG 459  
QY 183 AACAGCGCCTTCCTTCTACTACAGATFCAGCGTGGAGCAGTCCATCAGTCTCCCTACTC 242  
DB 460 AACAAAGCGCCTTCCTTCTACTACCAAAATAGTGTGGAACAATCTATCAGCCTCCTC 519  
QY 243 AATTTCATGTCGCGCTTACACAGTACGAGCGCGGAGGAGTTGACCGTGTTCAGNT 302  
DB 520 AACTTTATGATGTCGTCATATGACATGAGGCGCGGAGGCAAGTTCAGGATTTTCAGTT 579  
QY 303 AAAGCTATGTTAGCAACCATGTGTGTGGAATAATGCTGGACAAATTTGAGATACATTTTC 362  
DB 580 AAAGCTATGTTAGCAACCATGTGTGTGGAATAATGCTGGACAAATTTGAGATGATGTTTC 639  
QY 363 TCCAGATGTCAGATTCCTAATGGCTTAATGATGTTNGGAAAGCTTCAGCAGT 414  
DB 640 TCCAGATGTCAGATTCCTAATGGCTTAATGATGATTTAGCAAGTTTGACCAAGT 691

RESULT 2  
ID AAS73627  
XX

AC AAS73627;  
XX  
XX 13-FEB-2002 (first entry)  
XX  
XX DNA encoding novel human diagnostic protein #9431.  
XX  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
XX Homo sapiens.  
OS  
XX WO200175067-A2.  
PN  
XX 11-OCT-2001.  
PD  
XX 30-MAR-2001; 2001WO-US08631.  
XX  
XX 31-MAR-2000; 2000US-0540217.  
PR  
XX 23-AUG-2000; 2000US-0649167.  
XX  
XX (HYSE-) HYSEQ INC.  
PA  
XX Drmanac RT, Liu C, Tang YT;  
PI  
XX WPI; 2001-639362/73.  
DR  
XX P-PSDB; ABG09440.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
XX  
PS Claim 1; SEQ ID No 9431; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human  
CC diagnostic coding sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 2518 BP; 717 A; 627 C; 588 G; 586 T; 0 other;

Query Match 44.2%; Score 183; DB 23; Length 2518;  
Best Local Similarity 68.1%; Pred. No. 2e-52;  
Matches 280; Conservative 0; Mismatches 129; Indels 2; Gaps 2;  
QY 4 ACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTGT 63  
DB 164 ACCTACAGACAGCATGCAAGCT-TAGGTTTGTTCAGAGAATAATGCAATTTTCACCTGT 222  
QY 64 TCATATCTGGATATGATCGAGCTTCCGAGACAAATGGCCTTAACAGCTGGACACAG 123  
DB 223 GGACATATGGATATGATCGAGCATTCGGGAATAATGCTCTGAAACCTGGACCCAAA 282  
QY 124 CACGAGATCAGCGGTGTNCCGCTGGAGCGTCATCTCGTCCATCTACTATCAGTTGA 183  
DB 283 CACTGAACCTCA-ACGTGTCGCCCTTAGAGGCTGTCTCTCCACTATTTTACCAGCTCA 341



XX Claim 1; Page 115; 189pp; English.  
XX The invention relates to an isolated polynucleotide comprising a sequence  
CC selected from 183 human DNA sequences (appearing as ABK70130-ABK70312),  
CC or their fragments, homologues, variants or complements and their encoded  
CC polypeptides. Also included are an expression vector comprising the  
CC polynucleotide operably linked to an expression control sequence; a host  
CC cell transformed or transfected with an expression vector of; an isolated  
CC antibody, or its antigen-binding fragment that specifically binds to the  
CC polypeptide; a method for detecting the presence of a cancer in a  
CC patient; a fusion protein comprising at least the polypeptide; an  
CC oligonucleotide that hybridises to the polynucleotide under moderately  
CC stringent conditions; a method for stimulating and/or expanding T cells  
CC specific for a tumour protein; an isolated T cell population comprising T  
CC cells prepared from the method of above; a composition comprising a first  
CC component consisting of carriers and immunostimulants, and a second  
CC component selected from the polynucleotides, proteins, antibodies, fusion  
CC proteins, T cell populations and antigen presenting cells expressing the  
CC polypeptide; methods for stimulating an immune response or treating  
CC cancer in a patient by administering the composition and diagnostic kits  
CC comprising at least one of the oligonucleotide of, or an antibody and a  
CC detection reagent consisting of a reporter group. The polypeptides and  
CC polynucleotides are useful as vaccines for the treatment or prevention of  
CC lung cancer, and for diagnosis and monitoring of such cancer. The  
CC polynucleotide, polypeptide and antigen presenting cells can be  
CC used to stimulate or expand T cells specific for a tumorous protein.  
CC The polynucleotides may be used as probes or primers for nucleic acid  
CC hybridisation, and in the preparation of ribozyme molecules for  
CC inhibiting expression of tumour polypeptides and proteins in tumour  
CC cells. The present sequence is one of the 183 lung cancer associated  
CC polynucleotides.  
XX

SQ Sequence 512 BP; 127 A; 135 C; 141 G; 109 T; 0 other;

Query Match 43.7%; Score 181; DB 24; Length 512;  
Best Local Similarity 87.0%; Pred. No. 4e-52;  
Matches 221; Conservative 0; Mismatches 31; Indels 2; Gaps 2;  
QY 3 AACGTACAGACAGCTTGCAGTTACCGATTGTTACAGAGCGATGCAACCTTCATCTTG 62  
DB 260 AACTTACAGACAGCTGCGAATTA-CGATTGTACAAAAGCATGCAACCTTCATCTTG 318  
QY 63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAAGTGGCTTAACAGCTGGACCACA 122  
DB 319 TTGATATCTGGAACATGATTGAAGCCTTCCGAGACAAGTGGCTTAACAGCTGGACCACA 378  
QY 123 GCACGAGATCAGGGGTGTNCCGCTGGAGACCGCTCATCTCGTCCATCTACTATCATGTTG 182  
DB 379 CCACCGAGATCA-GTGTGTCCCGCTCGAAACTGTCTCTCCATCTACTATCATGTTG 437  
QY 183 ACAAGCGCTTCTTCTACTCACAGATCAGCGTGGAGAGTCATCATGCTCTCTACTC 242  
DB 438 ACAAGCGCTTCTTCTACTCACCAAAATAGTGTGGAACAATCTATCAGCGCTCTCTC 497  
QY 243 AATTTCATGCTGC 256  
DB 498 AACTTTATGATTGC 511

RESULT 5  
ABL14199  
ID ABL14199 standard; cDNA; 3045 BP.  
XX ABL14199;  
XX ABL14199;  
XX 26-MAR-2002 (first entry)  
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 37079.  
XX Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical; gene; ss.  
XX

OS Drosophila melanogaster.  
XX WO200171042-A2.  
XX 27-SEP-2001.  
XX 23-MAR-2001; 2001WO-US09231.  
XX 23-MAR-2000; 2000US-191637P.  
PR 11-JUL-2000; 2000US-0614150.  
XX (PEKE ) PE CORP NY.  
XX Venter JC, Adams M, Li PWD, Myers EW;  
PI WPI; 2001-658860/75.  
XX P-PSDB; ABB70096.  
DR New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX Claim 1; SEQ ID NO 37079; 21pp + Sequence Listing; English.  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB57737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX Sequence 3045 BP; 721 A; 882 C; 846 G; 596 T; 0 other;

Query Match 36.4%; Score 150.8; DB 23; Length 3045;  
Best Local Similarity 63.2%; Pred. No. 3.3e-41;  
Matches 244; Conservative 0; Mismatches 141; Indels 1; Gaps 1;  
QY 29 CGATTGTTACAGAGCGATGCAACCTTCATCTGTTGATATCTGATATGATCGAAGCT 88  
DB 602 CGCTATATACAGAGTCCACAATCTGCACCTGGTGACATATGGAATGATCGAGGCC 661  
QY 89 TTCCGAGACAATGGCTTTAACACGCTGGACCACACGACGAGATCAGCGGTGTCGGCT 148  
DB 662 TTTCGCGAAGTGGCTGAACACCTGGAGCGCGAGAGGTGA-GCGTGGCCAGGCT 720  
QY 149 GGAGACCGTCACTCGTCCATCTACTATCATGTTGACAGCGCTTCCTTCTACTCACA 208  
DB 721 GGAACCCCTGCTCTCTCTCTATCATAACTCAACAGCGCTGCCACCGCTCAGCA 780  
QY 209 GATCAGCGTGGAGAGTCCATCAGTCTCTACTCAATTTTCATGTCGCGCGCTACGAG 268  
DB 781 GGTGCTGTGACTCGAAGCGGGTCTGCTGCTCACTGCTGCTGGCGGTACACAG 840  
QY 269 TGAGGGCCGAGCGANGTTGACCGTGTGTTTCAGNTAAAGCTATGTTAGCAACCATGTG 328  
DB 841 TGATAACTCGGGCAAGATACCGGTTTTCTCCATCAAAGTGGCCCTGGCCACTATGTGTC 900  
QY 329 TGGAAAATGCTGGACAAATGAGATACATTTCTCCAGATGTCAGATTCATGGCTT 388  
DB 901 TCGCAAGCTGGTGACAACTAAGATACATATTTCTCGCAGATCTCGGATGGCGTGGACA 960  
QY 389 AATGATGTTNGGAAAGCTTGACCACT 414  
DB 961 GCTGTGCTCGAAGCTGGCGGAGT 986

RESULT 6  
ABL14198/c





```

AC ABL87347;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human ovarian cancer related cDNA clone SEQ ID NO:10325.
XX
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200192581-A2.
XX
PD 06-DEC-2001.
XX
PF 29-MAY-2001; 2001WO-US17756.
XX
PR 26-MAY-2000; 2000US-207484P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Algate PA, Harlocker SL, Jones R;
XX
WPI; 2002-122075/16.
XX
CC The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (III) of an ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.
XX
SQ Sequence 689 BP; 201 A; 156 C; 161 G; 162 T; 9 other;
Query Match 18.9%; Score 78.4; DB 24; Length 689;
Best Local Similarity 83.3%; Pred. No. 1.8e-16;
Matches 85; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

QY 313 TAGCAACCATGTGTGTGGAAAATGCTGGACAAATGAGATACATATTTCTCCAGATGT 372
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 688 TAGCANCATGTGTGTGGAAAATGCTGGACAAATGAGATATATGTTCTCCAGATGT 629
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 373 CAGATTCATGGCTTAATGATGTGTTGGAAGCTTGACCATG 414
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 628 CAGATTCATGGCTTAATGATATATTTAGCAAGTGTGACCATG 587
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 9
AAC01269
ID AAC01269 standard; cDNA; 458 BP.
XX
AC AAC01269;

```

---

```

XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein 5' EST, SEQ ID NO: 1267.
XX
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST ) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
WPI; 2000-500381/45.
XX
P-PSDB; AAG01263.
XX
CC New nucleic acid that is a 5' expressed sequence tag (5' EST) for
CC obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
CC diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
Claim 1; SEQ ID 1267; 71pp + CD-ROM; English.
XX
CC The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 458 BP; 145 A; 77 C; 114 G; 121 T; 1 other;
Query Match 17.2%; Score 71.2; DB 21; Length 458;
Best Local Similarity 69.7%; Pred. No. 4.4e-14;
Matches 122; Conservative 1; Mismatches 50; Indels 2; Gaps 2;

QY 4 ACGTACAGAACAGCTTGCAAGTTACCGATTGTGACAGAGCGATGCAACCTTCATCTTGT 63
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 286 ACCTACAGAACAGCATGCAAGCT-TAGGTTTGTTCAGAGAAATGCAATTTGCACCTGT 344
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 64 TGTATCTGGAATATGATGGAAGCTTCCGAGACATGCGCTTAACACGCTGGACCACAG 123
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 345 GGACATATGGAATGTGATGAGACGATTTGGGGAAATGCTCTGAGAACCTGGACCCAAA 404
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 124 CAGCGAGATCAGCGGTGTCGCCGCTGGAGACGTCATCTCGTCCATCTACTATCA 178
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 405 CACTGAACTCA-ACGTCGCCGCTTAGAGCTGTGCTCTCCACTATTTTTTACCA 458
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

RESULT 10
AAS31142
ID AAS31142 standard; cDNA; 5106 BP.
XX
AC AAS31142;
XX
DT 04-DEC-2001 (first entry)
XX

```

DE Human diagnostic and therapeutic polynucleotide (DITHP) #157.  
XX Human; receptor; diagnostic; therapeutic; gene therapy; vaccine;  
KW cell proliferative disorder; Crohn's disease; lymphoma; leukaemia;  
KW acquired immune deficiency syndrome; AIDS; autoimmune disorder;  
KW respiratory disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200162927-A2.  
XX  
PD 30-AUG-2001.  
XX  
PF 21-FEB-2001; 2001WO-US06059.  
XX  
PR 24-FEB-2000; 2000US-0184693.  
PR 24-FEB-2000; 2000US-0184697.  
PR 24-FEB-2000; 2000US-0184698.  
PR 24-FEB-2000; 2000US-0184768.  
PR 4-FEB-2000; 2000US-0184769.  
PR 24-FEB-2000; 2000US-0184770.  
PR 24-FEB-2000; 2000US-0184771.  
PR 24-FEB-2000; 2000US-0184772.  
PR 24-FEB-2000; 2000US-0184773.  
PR 24-FEB-2000; 2000US-0184774.  
PR 24-FEB-2000; 2000US-0184776.  
PR 24-FEB-2000; 2000US-0184777.  
PR 24-FEB-2000; 2000US-0184797.  
PR 24-FEB-2000; 2000US-0184813.  
PR 24-FEB-2000; 2000US-0184837.  
PR 24-FEB-2000; 2000US-0184841.  
PR 24-FEB-2000; 2000US-0185213.  
PR 24-FEB-2000; 2000US-0185216.  
PR 12-MAY-2000; 2000US-0203785.  
PR 15-MAY-2000; 2000US-0204226.  
PR 16-MAY-2000; 2000US-0204525.  
PR 16-MAY-2000; 2000US-0204821.  
PR 16-MAY-2000; 2000US-0204908.  
PR 16-MAY-2000; 2000US-0204815.  
PR 17-MAY-2000; 2000US-0205232.  
PR 17-MAY-2000; 2000US-0204863.  
PR 17-MAY-2000; 2000US-0205221.  
PR 17-MAY-2000; 2000US-0205285.  
PR 17-MAY-2000; 2000US-0205286.  
PR 17-MAY-2000; 2000US-0205287.  
PR 17-MAY-2000; 2000US-0205323.  
PR 17-MAY-2000; 2000US-0205324.  
XX  
PI . INCYTE GENOMICS INC.  
XX  
PI Panzer SR, Spiro PA, Banville SC, Shah P, Chalup MS, Chang SC;  
PI Chen A, D'Sa SA, Anshey S, Dahl CR, Dam TC, Daniels SE;  
PI Dufour GE, Flores V, Fong WT, Greenawalt LB, Hillman JL, Jones AL;  
PI Liu TF, Roseberry AM, Rosen BH, Russo FD, Stockdreher TK, Daffo A;  
PI Wright RJ, Yap PE, Yu JY, Bradley DL, Bratcher SR, Chen W;  
PI Cohen HJ, Hodgson DM, Lincoln SE, Jackson S;  
XX  
DR WPI: 2001-502867/55.  
DR P-PSDB; RAU19571.  
XX  
XX Polynucleotides encoding diagnostic and therapeutic proteins, e.g.  
XX enzymes, hormones and receptors, useful in diagnostics and therapeutics  
PT  
PT  
PT  
XX  
PS Claim 1; Page 370-371; 522pp; English.  
XX  
CC The invention relates to polynucleotides (I) encoding diagnostic and  
CC therapeutic (DITHP) polypeptides (II), which include e.g. enzymes,  
CC and proteins involved in growth and development and receptors. (I) and  
CC (II) may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate DITHP expression. For example, (I) and  
CC (II) may be used to treat disorders associated with decreased polypeptide  
CC expression by rectifying mutations or deletions in a patient's genome,  
PT

CC that affect the activity of the DITHPs, by expressing inactive proteins  
CC or supplementing the patient's own production of them. (I) and (II)  
CC may be used to treat diseases, for example, cell proliferative disorder,  
CC Crohn's disease, acquired immune deficiency syndrome (AIDS), lymphoma,  
CC leukaemia, autoimmune disorders, and respiratory disorders. Additionally,  
CC (I) may be used to produce the DITHPs, by inserting the nucleic acids  
CC into a host cell and culturing the cell to express the protein. (I) and  
CC its complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids in  
CC samples, and therefore which patients may be in need of restorative  
CC therapy. (II) may also be used as antigens in the production of  
CC antibodies against DITHPs and in assays to identify modulators of DITHP  
CC expression and activity. The anti-DITHP antibodies and antagonists may  
CC also be used to down regulate expression and activity. The anti-DITHP  
CC antibodies may also be used as diagnostic agents for detecting the  
CC presence of DITHPs in samples (e.g. by enzyme linked immunosorbent  
CC assay (ELISA)). AAS30986-AAS31196 represent human diagnostic and  
CC therapeutic (DITHP) polynucleotides of the invention.  
XX  
SQ Sequence 5106 BP; 1248 A; 1387 C; 1268 G; 1203 T; 0 other;

Query Match 9.5%; Score 39.2; DB 22; Length 5106;  
Best Local Similarity 51.1%; Pred. No. 0.022;  
Matches 89; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 206 CCAGATCAGCGGAGCAGTCATCAGTCTCTACTCAATTTTCAGTGGCGCGCTACCA 265  
DB 1711 CCTGTCAATGTGCCACTCTGTGTGACATGAGCTCAATGGCTCCTCAATGTTTGA 1770  
QY 266 CAGTGAGGCGCGAGCANGTTGACCGTGTTTTCAGNTAAAGCTATGTAGCAACCATGTG 325  
DB 1771 TAGTGTGCGCAGCGGAAGATCGGGCATGTGCTTTTAAAGATGGCATGTGCTGTG 1830  
QY 326 TGGTGGAAAAATGCTGGACAATTCAGATACATTTTCTCCAGATGTTCAGATTC 379  
DB 1831 TGGCAGGAAGTGAAGAAAAACTTCAGTACCTCTTCAGCCCAAGTGGCAACTC 1884

RESULT 11

AAV18885  
ID AAV18885 standard; cDNA; 13815 BP.

AC AAV18885;

DT 17-AUG-1998 (first entry)

XX Mus musculus dystrophin cDNA.

XX PCR primer; phage vector; bacteriophage lambda; gene therapy;  
KW antisense RNA; ribozyme; delivery; vaccine; antigen; treatment;  
KW prevention; in vivo; in vitro; AIDS; cancer; cystic fibrosis;  
KW acquired immune deficiency syndrome; Gaucher's disease; haemophilia;  
KW post-translational modification; diagnostic reagents; production;  
XX dystrophin; ss.  
OS Mus musculus.

XX WO9805344-A1.

XX 12-FEB-1998.

PD 03-JUL-1997; 97WO-US12928.

PF 11-MAR-1997; 97US-0814859.

XX 05-AUG-1996; 96US-0693865.

XX (BGHM ) BRIGHAM & WOMENS HOSPITAL.

XX Dublin DB, Kupper TS, Sarkar SN;

XX WPI: 1998-145345/13.

XX Improved gene therapy with phage vectors unable to release genetic  
PT

PT material into cells - comprises surface ligand binding to  
PT cell-surface receptor, and transcribing nucleic acid in cell, useful  
XX for, e.g. treating AIDS  
PS Example 9; Pages 80-85; 110pp; English.  
XX The sequence is that of murine full length dystrophin cDNA which.  
CC was used in construction of a chimeric bacteriophage lambda vector.  
XX  
SQ Sequence 13815 BP; 4582 A; 2731 C; 3012 G; 3490 T; 0 other;  
Query Match 9.2%; Score 38.2; DB 19; Length 13815;  
Best Local Similarity 53.0%; Pred. No. 0.085;  
Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 240 CTCATTTTCATGTCGCCCTACGACAGTGGGCGGAGGAGGATCGCTGCTGTCT 299  
DB 9734 CTCAACTGGCTTCTCAATGTTATGATACCGGACGAAGTGGAGATACCTT 9793  
QY 300 GNTAAGCTATGTAGCAACCATGTGTGTTGGAATAATCTGGACAAATTTGAGATACATT 359  
DB 9794 TTTAAACCTGGGATCATTTCTCTGTGTAAGACACACTTGGAAAGACAAGTACAGATACCTT 9853  
QY 360 TTCTCCAGATGTCAGATTCCTCAATGGCTT 388  
DB 9854 TTCAAGCAAGTGGCAAGTTCAACTGGCTT 9882  
RESULT 12  
ABK81960  
ID ABK81960 standard; DNA; 13815 BP.  
XX  
AC ABK81960;  
XX  
DT 13-AUG-2002 (first entry)  
XX  
DE cDNA encoding mouse dystrophin.  
XX  
KW Mini-dystrophin peptide; spectrin-like repeat domain; muscle disease;  
KW Duchenne's muscular dystrophy; DMD; dystrophin; mouse; gene; ds.  
XX  
OS Mus sp.  
XX  
PN WO200229056-A2.  
XX  
PD 11-APR-2002.  
XX  
PF 04-OCT-2001; 2001WO-US31126.  
XX  
XV 06-OCT-2000; 2000US-238848P.  
PA (UNMI ) UNIV MICHIGAN.  
XX  
PI Chamberlain JS, Harper SQ;  
XX  
DR WPI; 2002-435334/46.  
XX  
XX A composition for preparing therapeutic drugs, has a mini-dystrophin  
PT peptide comprising a specific number of spectrin-like repeat domains,  
PT or a nucleic acid sequence encoding the mini-dystrophin peptide -  
XX  
PS Example 1; Fig 2; 145pp; English.  
XX  
CC The invention describes a composition comprising a mini-dystrophin  
CC peptide comprising a spectrin-like repeat domain, where the domain  
CC comprises n spectrin-like repeats, and contains no more than n  
CC spectrin-like repeats, where n is an even number between 4-24, or a  
CC nucleic acid encoding a mini-dystrophin peptide. The mini-dystrophin  
CC peptide or the polynucleotide encoding it is useful as a medicament,  
CC for preparing a drug for therapeutic application and in the preparation  
CC of a composition for treatment of muscle disease, e.g. Duchenne's  
CC muscular dystrophy (DMD). This sequence encodes mouse dystrophin used  
CC in the creation of the mini-dystrophin peptides of the invention.

XX  
SQ Sequence 13815 BP; 4577 A; 2719 C; 3025 G; 3494 T; 0 other;  
Query Match 9.2%; Score 38.2; DB 24; Length 13815;  
Best Local Similarity 53.0%; Pred. No. 0.085;  
Matches 79; Conservative 0; Mismatches 70; Indels 0; Gaps 0;  
QY 240 CTCATTTTCATGTCGCCCTACGACAGTGGGCGGAGGAGGATCGCTGCTGTCT 299  
DB 9734 CTCAACTGGCTTCTCAATGTTATGATACCGGACGAAGTGGAGATACCTT 9793  
QY 300 GNTAAGCTATGTAGCAACCATGTGTGTTGGAATAATCTGGACAAATTTGAGATACATT 359  
DB 9794 TTTAAACCTGGGATCATTTCTCTGTGTAAGACACACTTGGAAAGACAAGTACAGATACCTT 9853  
QY 360 TTCTCCAGATGTCAGATTCCTCAATGGCTT 388  
DB 9854 TTCAAGCAAGTGGCAAGTTCAACTGGCTT 9882  
RESULT 13  
ABI99799  
ID ABI99799 standard; cDNA; 13815 BP.  
XX  
AC ABI99799;  
XX  
DT 07-MAR-2002 (first entry)  
XX  
DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:904.  
XX  
KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;  
KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.  
XX  
OS Mus musculus.  
XX  
PN WO200188188-A2.  
XX  
PD 22-NOV-2001.  
XX  
PF 18-MAY-2001; 2001WO-JP04192.  
XX  
PR 18-MAY-2000; 2000JP-0145977.  
XX  
PA (UYN1-) UNIV NIHON SCHOOL JURIDICAL PERSON.  
XX  
PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;  
XX  
DR WPI; 2002-034733/04.  
XX  
PT Examining the ischemic condition (e.g. occlusive ischemia) by measuring  
PT expression levels of particular genes defined in the specification or  
PT by determining the expression profile of a gene group comprising these  
PT genes -  
XX  
PS Claim 2; Page 2281-2289; 2690pp; English.  
XX  
CC The present invention describes a method for examining ischaemic  
CC conditions, comprising measuring the expression levels of particular  
CC genes (I) in a test sample or determining the expression profile of a  
CC gene group in the sample comprising genes selected from (I). The method  
CC is useful for examining the ischaemic condition (e.g. compressive  
CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring  
CC expression levels of particular genes (ABI99202 to ABI99912, encoding  
CC the protein sequences in ABB57020 to ABB57374) or by determining the  
CC expression profile of a gene group comprising these genes. The  
CC expression levels or expression profiles produced by these genes are  
CC used as an indicator when screening for ischaemic condition-improving  
CC drugs or therapeutics for ischaemic diseases. ABI99913 and ABI99914  
CC represent PCR primers for a mouse ischaemic condition related sequence,  
CC which are used in the exemplification of the present invention.  
XX  
SQ Sequence 13815 BP; 4577 A; 2719 C; 3025 G; 3494 T; 0 other;



Db . 288 AGCACATTTTGGAGACAAGTACAGATACCTTTTCAAGCAAGTGGCAAGTTCAACAGGATT 347

Search completed: December 1, 2002, 16:24:34  
Job time : 290 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2002 Compugen Ltd.  
Run on: December 1, 2002, 16:20:13 ; Search time 53 seconds  
(without alignments)  
3008.260 Million cell updates/sec

Title: US-09-680-959-125  
Perfect score: 414  
Sequence: 1 ctaacgtacagaacagcttg.....gttnggaagcttgaccagt 414  
Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0  
341543 seqs, 192557720 residues  
To. number of hits satisfying chosen parameters: 683086  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications, NA:  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:  
5: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_NEW\_PUB.seq:  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:  
8: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:  
9: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:  
10: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:  
11: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:  
14: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

result No.	Score	Query Match	Length	ID	Description
1	318.4	76.9	2247	10	US-09-960-253-157
2	181	43.7	512	10	US-09-960-253-157
3	78.4	18.9	689	10	US-09-960-253-157
4	37.2	9.0	13957	10	US-09-867-701-10325
5	37.2	9.0	13957	10	US-09-867-701-10325
6	32.2	7.8	1173	9	US-09-782-378A-22
7	32	7.4	1040	9	US-09-880-107-2284
8	30.8	7.4	80959	9	US-09-938-842A-1992
9	30	7.2	1098	9	US-09-825-345-1
10	29.6	7.1	1927	9	US-09-815-242-7522
11	29.6	7.1	2734	9	US-09-764-868-66
12	29.6	7.1	2964	10	US-09-764-868-489
13	29.2	7.1	307	10	US-09-198-559-2
14	29.2	7.1	348	10	US-09-294-093B-1641
15	29.2	7.1	442	10	US-09-864-761-21546
16	28.8	7.0	2010	10	US-09-864-761-4808
17	28.6	6.9	357	10	US-09-887-576-301
18	28.6	6.9	402	10	US-09-878-574-2805
19	28.6	6.9	408	10	US-09-777-564-462
20	28.6	6.9	408	10	US-09-878-574-4089

ALIGNMENTS  
RESULT 1  
US-09-960-253-157  
; Sequence 157, Application US/09960253  
; Patent No. US20020123619A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Mohamath, Raodoh  
; APPLICANT: Lodes, Michael J.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: AND DIAGNOSIS OF LUNG CANCER  
; CURRENT APPLICATION NUMBER: 210121.556  
; NUMBER OF SEQ ID NOS: 187  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 157  
; LENGTH: 2247  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-960-253-157

QY	3	AACGTACAGAACAGCTTGCAGCTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG	76.9%;	Score 318.4;	DB 10;	Length 2247;
Db	282	AACGTACAGAACAGCTTGCAGCTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG	88.6%;	Pred. No. 4.2e-99;		
QY	63	TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA	45;	Indels	2;	Gaps
Db	341	TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA				
QY	123	GCACGGAGATCAGGCGTTCGAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA				
Db	401	GCACGGAGATCAGGCGTTCGAGCTTTCGAGACAAATGGCCTTACACGCTTGACACCA				
QY	183	AACAAGCGCTTCTTCTTACTACACGATCAGGTCGAGCTTCTTCTTCTTCTTCTTCTTCTT				
Db	460	AACAAGCGCTTCTTCTTACTACACGATCAGGTCGAGCTTCTTCTTCTTCTTCTTCTTCTT				
QY	243	AATTTCATGGTGGCGCTTACTACACGATCAGGTCGAGCTTCTTCTTCTTCTTCTTCTTCTT				

us-09-680-959-125.rnpb

ue Dec 3 12:27:59 2002

```

; SEQ ID NO 10325
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(689)
; OTHER INFORMATION: n = A,T,C or G
; US-09-867-701-10325

Query Match      18.9%; Score 78.4; DB 10; Length 689;
Best Local Similarity 83.3%; Pred. No. 5.6e-17; Indels 0; Gaps 0;
Matches 85; Conservative

QY 313 TAGCAACCATGCTGTGGTGGAAATGCTGGCAAAATTTGAGATACATTTTCTCCAGATGT 372
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 688 TAGCANCCATGCTGTGGTGGAAATGCTGGCAAAATTTGAGATACATTTTCTCCAGATGT 372
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 373 CAGATTCCAAATGGCTTAATGATGTTGGAAAGCTTGACCAAGT 414
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 CAGATTCCAAATGGCTTAATGATGTTGGAAAGCTTGACCAAGT 414

RESULT 4
US-09-782-378A-22
; Sequence 22, Application US/09782378A
; Patent No. US20020102731A1
; GENERAL INFORMATION:
; APPLICANT: Hearing, Patrick
; APPLICANT: Bahou, Wadie
; APPLICANT: Sandalon, Ziv
; APPLICANT: Gnatenko, Dmitri
; TITLE OF INVENTION: Adenoviral Vectors
; FILE REFERENCE: STONYB-04970
; CURRENT APPLICATION NUMBER: US/09/782,378A
; CURRENT FILING DATE: 2001-02-12
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-782-378A-22

Query Match      9.0%; Score 37.2; DB 10; Length 13957;
Best Local Similarity 50.0%; Pred. No. 0.05; Indels 0; Gaps 0;
Matches 90; Conservative

QY 209 GATCAGCGTGGAGCAGTCCATCAGTCTCTCTCTGCTGGATATGTTCTGAACCTGGCTGGAATGTTATGATAC 9771
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9712 GGTCAACGTCCTCTCTCTGCTGGATATGTTCTGAACCTGGCTGGAATGTTATGATAC 9771
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 269 TGAGGCGCCAGGCGANGTTGACCGTGTTCAGNTAAAGTATGTTAGCAACCAATGTGTG 328
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9772 GGGACGAACAGGAGGATCCGTCCTGTCTTTTAAACATGGCATCATTTCTCCAGATGTCAATTCATGCTT 388
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 329 TGGAAATGCTGGCAAAATTTGAGATACATTTCTCCAGATGTCAATTCATGCTT 9891
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 9832 AGCATTGGAAGACAGTACATCTTTTCAAGCAAGTGGCAGTTCAACAGGATT 9891
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-880-107-2284
; Sequence 2284, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer

```

```

; US-09-960-253-28
; Sequence 28, Application US/09960253
; Patent No. US20020123619A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Lodes, Michael J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.556
; CURRENT APPLICATION NUMBER: US/09/960,253
; CURRENT FILING DATE: 2001-09-20
; NUMBER OF SEQ ID NOS: 187
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 512
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-960-253-28

Query Match      43.7%; Score 181; DB 10; Length 512;
Best Local Similarity 87.0%; Pred. No. 2.7e-52; Indels 2; Gaps 2;
Matches 221; Conservative

QY 3 AAGTCACAGAACAGCTTGGCAAGTTCGAGACATGCGCTTAACAGCTGGACACA 122
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 260 AACTTACAGAACAGCTTGGCAAGTTCGAGACATGCGCTTAACAGCTGGACACA 122
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 53 TTGATATCTGGAATGATGATCGAAGCTTCCGAGACATGCGCTTAACAGCTGGACACA 378
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 319 TTGATATCTGGAATGATGATCGAAGCTTCCGAGACATGCGCTTAACAGCTGGACACA 378
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 123 GCACGGAGATAGGCGTGTGTCGCTGAGACCGTCACTCTGCTCATCTACTATCAGTTG 182
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 379 CCACCGAGATCA-GTGTGTCGCTGAGACCGTCACTCTGCTCATCTACTATCAGTTG 437
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 183 ACAAGCGCTTCTCTCTACTACAGATCAGCGTGGAGACGTCCTACTCTCTCTCTC 242
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 438 ACAAGCGCTTCTCTCTACTACAGATCAGCGTGGAGACGTCCTACTCTCTCTCTC 497
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 243 AATTTCATGTCGC 256
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 498 AACTTATGATTGC 511
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
US-09-867-701-10325/c
; Sequence 10325, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSPO for Windows Version 4.0

```







Patent No. US20020025930A1  
GENERAL INFORMATION:  
APPLICANT: Hillman, Jennifer L.  
APPLICANT: Corley, Neil C.  
APPLICANT: Shah, Purvi  
TITLE OF INVENTION: RAS-LIKE PROTEIN  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Incyte Pharmaceuticals, Inc.  
STREET: 3174 Porter Drive  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/198,559  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/846,790  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PF-0388 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2964 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: COLNUT16  
CLONE: 2791521  
US-09-198-559-2

Query Match 7.1%; Score 29.6; DB 10; Length 2964;  
Best Local Similarity 53.0%; Pred. No. 7.9;  
Matches 62; Conservative 0; Mismatches 55; Indels 0; Gaps 0;  
QY 58 TCTTGTGATATCTGATATGATGAGCTTTCCGAGACAATGGCCTTAACACGCTGGA 117  
DB 863 TCTCTTTTCATATACAGCAACGGGCAAGGTTTTTGGGTCTATCAGGAGATCAATCTGGG 804  
QY 118 CCACAGCAGGAGATCAGCGGTGCCCTCGAGACCGTCTATCTCGTCCATCTACT 174  
DB 803 TCCCTATGAGGACATAGGACGATGAGTCTTGAGTCTGGGACCCCAT 747

RESULT 13  
US-09-294-093B-1641  
Sequence 1641, Application US/09294093B  
Patent No. US20010051335A1  
GENERAL INFORMATION:  
APPLICANT: Lalgudi, Raghunath, V.  
APPLICANT: Ito, Laura, Y.  
APPLICANT: Sherman, Bradley, K.  
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL  
FILE REFERENCE: PL-0009 US  
CURRENT APPLICATION NUMBER: US/09/294,093B  
CURRENT FILING DATE: 1999-04-16  
PRIOR APPLICATION NUMBER: 60/082,567  
PRIOR FILING DATE: April 21, 1998  
NUMBER OF SEQ ID NOS: 6207

SOFTWARE: PERL Program  
SEQ ID NO 1641  
LENGTH: 307  
TYPE: DNA  
ORGANISM: Zea mays  
FEATURE:  
NAME/KEY: misc\_feature  
OTHER INFORMATION: Incyte ID No. US20010051335A1 700344740H1  
NAME/KEY: unsure  
LOCATION: 12-13, 30-31, 41, 43-44, 70, 79, 81, 105-106, 140, 153, 178, 192, 196,  
OTHER INFORMATION: a, t, c, g, or other  
US-09-294-093B-1641

Query Match 7.1%; Score 29.2; DB 10; Length 307;  
Best Local Similarity 51.8%; Pred. No. 2.7;  
Matches 58; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 139 TGTNCGCCTCGAGACCGTCATCTGTCATCTACTATCATCTAGTGAACAAGCGCTTCCTT 198  
DB 146 TATATCANGTTGACGGGACAGAGATCCATCTCTATAAGCGCTNAAAANTGCCCTCTG 205  
QY 199 CTACTCACCAGATCAGCGGTGGAGCAGTCATCAGTCTCTCTACTCAATTCAT 250  
DB 206 CGATTAAACAGTTCACCCAGGCTGGACCGCAACAGCTATNCAGCTGCTT 257

RESULT 14  
US-09-864-761-21546/C  
Sequence 21546, Application US/09864761  
Patent No. US20020048763A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharron G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
APPLICANT: Chen, Wensheng  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
FILE REFERENCE: Aecomica-X-1  
CURRENT APPLICATION NUMBER: US/09/864,761  
CURRENT FILING DATE: 2001-05-23  
PRIOR APPLICATION NUMBER: US 60/180,312  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: US 60/207,456  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: US 09/632,366  
PRIOR FILING DATE: 2000-08-03  
PRIOR APPLICATION NUMBER: GB 24263.6  
PRIOR FILING DATE: 2000-10-04  
PRIOR APPLICATION NUMBER: US 60/236,359  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: PCT/US01/00666  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00667  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00664  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00669  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00665  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00668  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00663  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00662  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00661  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: PCT/US01/00670  
PRIOR FILING DATE: 2001-01-30  
PRIOR APPLICATION NUMBER: US 60/234,687  
PRIOR FILING DATE: 2000-09-21  
PRIOR APPLICATION NUMBER: US 09/608,408



GenCore version 5.1.3  
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 1, 2002, 16:14:38 ; Search time 1993 Seconds  
(without alignments)  
3381.206 Million cell updates/sec

Title: US-09-680-959-125

Perfect score: 414

Sequence: 1 ctaacacacagacagcttg.....gttnggaagcttgaccagt 414

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

1 . number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em.estba:\*

2: em.esthum:\*

3: em.estlin:\*

4: em.estmu:\*

5: em.estov:\*

6: em.estpl:\*

7: em.estro:\*

8: em.htc:\*

9: gb.est1:\*

10: gb.est2:\*

11: gb.htc:\*

12: gb.est3:\*

13: gb.est4:\*

14: gb.est5:\*

15: em.estfun:\*

16: em.estom:\*

17: gb.gss:\*

18: em.gss\_hum:\*

19: em.gss\_inv:\*

20: em.gss\_pln:\*

21: em.gss\_vrt:\*

22: em.gss\_fun:\*

23: em.gss\_mam:\*

24: em.gss\_mus:\*

25: em.gss\_other:\*

26: em.gss\_pro:\*

27: em.gss\_rod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	382.4	92.4	505	10	BE650744 UI-M-BH3-
2	363	87.7	517	9	AI325959 mq63d07.x
3	348.2	84.1	958	12	BG865074 602784226
4	330.8	79.9	676	10	BB642867 BB642867
5	328	79.2	461	10	AW496140 UP47e06.y
6	318.4	76.9	791	12	BG700652 602682269

7	318.4	76.9	880	14	BQ423333
8	318.4	76.9	932	14	BQ651659
9	318.4	76.9	967	9	AL524324
10	318	76.8	996	9	AL555818
11	317	76.6	945	13	BI854854
12	311.4	75.2	647	10	BB654888
13	307	74.2	590	12	BG078688
14	303	73.2	600	12	BG805399
15	299.4	72.3	830	12	BG469017
16	299	72.2	600	12	BG804143
17	299	72.2	622	10	BE500997
18	295	71.3	448	10	AW495863
19	293	70.8	673	10	BB640953
20	290.8	70.2	823	9	AL527400
21	290.4	70.1	571	10	BE552374
22	279.6	67.5	760	9	AL538124
23	277.2	67.0	915	12	BE779121
24	273.6	66.1	595	9	AJ454452
25	270.4	65.3	712	10	AW327909
26	261.6	63.2	954	12	BG115265
27	260.8	63.0	584	13	BI339696
28	255	61.6	654	13	BJ065983
29	245	59.2	926	9	AL530827
30	243	58.7	878	12	BG702733
31	242	58.5	419	12	BE947648
32	241	58.2	457	12	BE950528
33	240.2	58.0	422	10	AW491035
34	212	51.2	381	12	BF721737
35	210.4	50.8	541	9	AL636722
36	210	50.7	352	10	BE113690
37	194.6	47.0	390	10	BE379323
38	194	46.9	464	10	BE482496
39	194	46.9	574	12	BG689246
40	190.8	46.1	396	9	AI574752
41	190	45.9	1061	14	BN918263
42	186.2	45.0	918	14	BQ891601
43	186.2	45.0	1608	11	AK003829
44	186.2	45.0	3861	11	AK019477
45	184.2	44.5	701	13	BI546456

#### ALIGNMENTS

RESULT 1  
BE650744/C  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
COMMENT

BE650744  
UI-M-BH3-asr-h-05-0-UI-r1 NIH\_BMAP\_M\_S4 Mus musculus cDNA clone  
UI-M-BH3-asr-h-05-0-UI 5', mRNA sequence.  
BE650744.1 GI:9976616  
EST.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 505)  
Ronald M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
9704477  
Contact: Chin, H  
National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890  
Email: mstetnail.nih.gov  
cDNA Library Preparation: M.B. Soares Lab Clone distribution:  
Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It  
should be noted that Bento Soares is generating a small number of  
additional specialized non-redundant arrays of BMAP cDNAs whose

880	14	BQ423333	AGENCOURT
932	14	BQ651659	AGENCOURT
967	9	AL524324	AL524324
996	9	AL555818	AL555818
945	13	BI854854	603381829
647	10	BB654888	BB654888
590	12	BG078688	H3030G07-
600	12	BG805399	0692-19 M
830	12	BG469017	602510757
600	12	BG804143	0252-09 M
622	10	BE500997	7a34c05.x
448	10	AW495863	up47e06.x
673	10	BB640953	BB640953
823	9	AL527400	AL527400
571	10	BE552374	hy06g10.x
760	9	AL538124	AL538124
915	12	BE779121	601464731
595	9	AJ454452	AJ454452
712	10	AW327909	dr02d11.x
954	12	BG115265	602316195
584	13	BI339696	364749 MA
654	13	BJ065983	BJ065983
926	9	AL530827	AL530827
878	12	BG702733	602684286
419	12	BE947648	BE947648
457	12	BE950528	UI-M-BH3-
422	10	AW491035	UI-M-BH3-
381	12	BF721737	mab22d06.
541	9	AL636722	AL636722
352	10	BE113690	UI-R-BJ1-
390	10	BE379323	MRO-HT024
464	10	BE482496	168310 BA
574	12	BG689246	BG689246
396	9	AI574752	UI-R-GO-U
1061	14	BN918263	BN918263
918	14	BQ891601	AGENCOURT
1608	11	AK003829	Mus muscu
3861	11	AK019477	Mus muscu
701	13	BI546456	603188757



```

Db 249 GCACGGAGATCA-GCGTGTCCCGCTGGAGACCGTCACTCGTCCATCTACTATCAGTTG 307
|||||
QY 183 AACAAAGCGCCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 242
|||||
Db 308 AACAAAGCGCCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 367
|||||
QY 243 AATTTCATGTGCGCGCCTACGACAGTACGAGGCGCGAGGCGGAGTGGACCGTGTGTTTTCAGNT 302
|||||
Db 368 AATTTCATGTGCGCGCCTACGACAGTACGAGGCGCGAGGCGGAGTGGACCGTGTGTTTTCAGTT 427
|||||
QY 303 AAAGCTATGTAGCAACCATGTGTGGTGGAAAAATGCTGACAAAAATGAGATACATTTTC 362
|||||
Db 428 AAAGCTATGTAGCAACCATGTGTGGTGGAAAAATGCTGACAAAAATGAGATACATTTTC 487
|||||
QY 363 TCCAGATGTCAGATTCCAATGCTTAATG 392
|||||
Db 488 TCCAGATGTCAGATTCCAATGCTTAATG 517
|||||

R.L._J 3
BG865074
LOCUS 602784226F1 NCI_CGAP_SG2 Mus musculus cDNA clone IMAGE:4910166 5',
DEFINITION mRNA sequence.
ACCESSION BG865074
VERSION BG865074.1 GI:14215612
KEYWORDS EST.
SOURCE house mouse,
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 958)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10810 row: d column: 07
High quality sequence stop: 734.

F' RES Location/Qualifiers
Source l . 958
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4910166"
/clone_lib="NCI_CGAP_SG2"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: salivary gland; Vector: pCMV-SPORT6; Site: 1.
Noti; Site: 2: Salt; Cloned unidirectionally. Primer: Oligo
dt. Average insert size 1.3 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 247 a 264 c 235 g 212 t
ORIGIN

Query Match 84.1%; Score 348.2; DB 12; Length 958;
Best Local Similarity 95.9%; Pred. No. 1.1e-96;
Matches 398; Conservative 0; Mismatches 12; Indels 5; Gaps 4;

QY 3 ACGTACAGACAGCTTGCAGTTACCGATTGTTACAGAGCGATGCAACCTTCATCTTG 62
|||||
Db 292 AACGTACAGACAGCTTGCAGTTA-CGATTGTACAGAGCGATGCAACCTTCATCTTG 350
|||||
QY 63 TTGATATCTGGATATGATCGAAGCTTCCGACAGCAATGGCTTACACGCTGGACCACA 122
|||||
Db 351 TTGATATCTGGATATGATCGAAGCTTCCGACAGCAATGGCTTACACGCTGGACCACA 410
|||||

```

```

QY 123 GCACGAGATCAGCGCTGCCCTGGAGACCGTCACTCGTCCATCTACTATCAGTTG 182
|||||
Db 411 GCACGAGATCA-GCGTGTCCCGCTGGAGACCGTCACTCGTCCATCTACTATCAGTTG 469
|||||
QY 183 AACAAAGCGCCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 242
|||||
Db 470 AACAAAGCGCCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 529
|||||
QY 243 AATTTCATGTGCGCGCCTACGACAGTACGAGGCGCGAGGCG-ANGTTGACCGTGTGTTTCAGN 301
|||||
Db 530 AATTTCATGTGCGCGCCTACGACAGTACGAGGCGCGAGGCGAAAGTTGACCGTGTGTTTCAGT 589
|||||
QY 302 TAAAGCTATGTAGCAACCATGTGTGGTGGAAAAA--TGCTGGACAAAATGAGATACATT 359
|||||
Db 590 TAAAGCTATGTAGCAACCATGTGTGGTGGAAAAAATGCTGGACAAAATGAGATACATT 649
|||||
QY 360 TTCTCCAGATGTCAGATTCCAATGCTTAATGTTTNGGAAAGCTTGACCAGT 414
|||||
Db 650 TTCTCCAGATGTCAGATTCCAATGCTTACTGTCTGTGGAAGCTTGACCAGT 704
|||||

RESULT 4
BB642867
LOCUS BB642867 RIKEN full-length enriched, adult retina Mus musculus cDNA
DEFINITION clone A930104E03 5', mRNA sequence.
ACCESSION BB642867
VERSION BB642867
KEYWORDS EST.
SOURCE house mouse,
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 676)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okada
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished (2001)
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y., and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.

```

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (<http://genome.9sc.riken.go.jp>) for further details.  
e mouse tissues.

## FEATURES

Location/Qualifiers

1. 676

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="A930104E03"

/clone\_lib="RIKEN full-length enriched, adult retina"

/tissue\_type="retina"

/dev\_stage="adult"

/lab\_host="DH10B"

/note="Site\_1: Sall; Site\_2: BamHI; cDNA library was

prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5',

GAGAGAGAGGATCCAGAGCTCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5'GAGAGAGATTCGAGTTAATAATATCCCTCCCTCCCTCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda PLC I. -Retina RNA was provided by Stefano Gustincich, Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA02115, USA, whose assistance we gratefully acknowledge."

BASE COUNT 150 a 180 c 200 g 144 t 2 others

ORIGIN

Query Match 79.9%; Score 330.8; DB 10; Length 676;

Best Local Similarity 97.2%; Pred. No. 2.1e-91;

Matches 376; Conservative 0; Mismatches 7; Indels 4; Gaps 4;

QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62

|||||

Db 280 AACGTACAGACAGCTTGCAAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 338

QY 63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAAATGGCCTTAACACGCTGGACCACA 122

|||||

Db 339 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAAATGGCCTTAACACGCTGGACCACA 398

123 GCACGGAGATCAGCGCTGTGCCCTGGAGACCGTCCATCTCGTCCATCTACTATCAGTTG 182

|||||

Db 399 GCACGGAGATCA-CGGTGTCCCGCTGGAGACCGTCCATCTCGTCCATCTACTATCAGTTG 457

QY 183 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCATCTCTCTACTTC 242

|||||

Db 458 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCATCTCTCTACTTC 517

QY 243 AATTTTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 302

|||||

Db 518 AATNTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 577

QY 303 AAAGCTATGTAGCAACCATGTGTGGTGGAAAATGCTGGACAAATTTGAGATACATTTTC 362

|||||

Db 578 AAAGCTATGTAGCAACCATGTGTGGTGGAAAATGCTGGACAAATTTGAGATACATTTTC 636

QY 363 TCCAGATGTCAGATTCCTAATGGCTTTA 389

|||||

Db 637 T-CCAGATGTCAGATTCCTAATGGCTTTA 662

RESULT 5

AW496140/c

LOCUS

AW496140

461 bp

mRNA

linear

EST 24-FEB-2000

## DEFINITION

up47e06.y1 Soares\_mouse\_NMGB\_bcell Mus musculus cDNA clone  
IMAGE:2749282 5' similar to TR:070585 070585 DYSTROBIN, BETA ;,  
mRNA sequence.

ACCESSION AW496140.1 GI:7066421

VERSION AW496140.1

KEYWORDS EST.

SOURCE house mouse.

## ORGANISM

Mus musculus

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE 1 (bases 1 to 461)

AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Other ESTs: up47e06.x1

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.

MGI:1035654

Possible reversed clone: similarity on wrong strand

Seq primer: -40RP from Gibco.

## FEATURES

source

1. 461

/organism="Mus musculus"

/db\_xref="taxon:10090"

/clone="IMAGE:2749282"

/clone\_lib="Soares\_mouse\_NMGB\_bcell"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: germinal B-cell; Vector: pT7T3D-Pac

(Pharmacia) with a modified polylinker; Site\_1: Not I;

Site\_2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dr) primer [5,

TGTTACCAATCTGAAGTGGAGCGGCCCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT7T3 vector.

Library is normalized; constructed by Bento Soares and

M.Fatima Bonaldo."

BASE COUNT 107 a 117 c 109 g 128 t

Query Match 79.2%; Score 328; DB 10; Length 461;

Best Local Similarity 98.6%; Pred. No. 1.2e-90;

Matches 350; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62

|||||

Db 353 AACGTACAGACAGCTTGCAAGTTA-CGATTTGTACAGAGCGATGCAACCTTCATCTTG 295

QY 63 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAAATGGCCTTAACACGCTGGACCACA 122

|||||

Db 294 TTGATATCTGGAATATGATCGAAGCTTCCGAGACAAATGGCCTTAACACGCTGGACCACA 235

QY 123 GCACGGAGATCAGCGCTGTGCCCTGGAGACCGTCCATCTCGTCCATCTACTATCAGTTG 182

|||||

Db 234 GCACGGAGATCA-CGGTGTCCCGCTGGAGACCGTCCATCTCGTCCATCTACTATCAGTTG 176

QY 183 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCATCTCTCTACTTC 242

|||||

Db 175 AACAGCGCTTCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCATCTCTCTACTTC 116

QY 243 AATTTTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 302

|||||

Db 115 AATTTTCATGGTCCGCCCTACGACAGTGGAGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGG 56

QY 303 AAAGCTATGTAGCAACCATGTGTGTGGAAAATGCTGGACAAATTTGAGATACATA 357

|||||

Db 55 AAAGCTATGTAGCAACCATGTGTGTGGAAAATGCTGGACAAATTTGAGATACATA 1

RESULT 6

Db 573 AAGAGTATGTTAGCAACCATGCTGCTGCAAAAAATGCTGGACAATTTGACATATGTTTC 633

Qy 363 TCCAGAGTCAGATCCCAATGGCTTAATGATGTTNGGAAAGCTTGACCACT 414

Db 633 TCCAGAGTCAGATCCCAATGGCTTAATGATGTTAGCAAGTTGACCACT 684

RESULT 7

BQ423333

LOCUS BQ423333 880 bp mRNA linear EST 23-MAY-2003

DEFINITION AGENCOURT\_7846069 NIH\_MGC\_72 Homo sapiens cDNA clone IMAGE:6064183

ACCESSION BQ423333 5', mRNA sequence.

VERSION BQ423333.1 GI:21118648

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1 (bases 1 to 880)

TITLE NTH-MGC http://mgc.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC/DCTD/DTP

cDNA Library Preparation: Life Technologies, Inc.

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LLAM13338 row: h column: 08

High quality sequence stop: 714.

Location/Qualifiers

1. 880

Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="IMAGE:6064183"

/clone\_lib="NIH\_MGC\_72"

/tissue\_type="melanotic melanoma"

/lab\_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pCMV-Sport6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 2 kb. Library constructed by Life Technologies."

BASE COUNT 237 a 211 c 204 g 227 t 1 others

ORIGIN

Query Match 76.98; Score 318.4; DB 14; Length 880;

Best Local Similarity 88.68; Pred. No. 1.7e-87;

Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2

Qy 3 AAGTCAGACAACAGCTTGCAAGTTACCGATTGTGTACAGAAGCGATGCAACCTTCATCTTG 62

Db 168 AACTTACAGAACAGCTTGCAAAATTA-CGATTTGTACAAACAGATGCAACCTTCATCTTG 226

Qy 63 TTGATATCTGGAATATGATCGAAGCTTTCGAGACAATGGCTTTAAACAGCTGGACCACA 122

Db 227 TTGATATCTGGAACATGATTGAAGCCCTCCGAGACAATGGCTTTAAATACACTGGACCATA 286

Qy 123 GCAGGAGATCAGCGGTGTCGCCCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182

Db 287 CCACCGAGATCA-GTGTGTGCCCGCTCGAAACTGTCTATCTCTCCATCTACTATCAGTTG 345

Qy 183 AACAGCGCCTTCCTTCTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 242

Db 346 AACAAGCGCCTTCCTTCTACTCACCAGATGAGTGGAGCAACTTATCAGCGCTCTCCTC 405

Qy 243 AATTTCATGTCGCGCCTTACGAGTGAAGGCGGAGGCGANGTTGACCGTGTTCAGNT 302

Db 406 AACTTTATGATTGCTGCATATGACAGTGAAGGCGGAGGCGGAGTTGACGGTATTTCAGTT 465



QY 303 AAGCATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362  
 |||||  
 Db 466 AAGCATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 525  
 |||||  
 QY 363 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTNGGAAAGCTTGACCAAGT 414  
 |||||  
 Db 526 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTAGCAAGTTTGACCAAGT 577  
 |||||

RESULT 8  
 BQ651659 932 bp mRNA linear EST 15-JUL-2002  
 LOCUS AGENCOURT\_8342332 NIH\_MGC\_100 Homo sapiens cDNA clone IMAGE:6268278  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BQ651659  
 VERSION BQ651659.1 GI:21775831  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 932)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
 Tissue Procurement: CGAP (Stanford)  
 cDNA Library Preparation: Rubin Laboratory  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 Cloning Strategy: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
 Plate: LLCM2439 row: p column: 07  
 High quality sequence stop: 733.

FEATURES  
 Location/Qualifiers  
 1..932  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6268278"  
 /clone\_lib="NIH\_MGC\_100"  
 /tissue\_type="hepatocellular carcinoma, cell line"  
 /lab\_host="DH10B (phage-resistant)"  
 /note="Organ: liver; Vector: pOTB7; Site:1: xhoI; Site:2:  
 EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
 into EcoRI/XhoI sites using the following 5' adaptor:  
 GGCAGGAG(G). Size-selected >500bp for average insert size  
 1.8kb. Library constructed by Ling Hong in the laboratory  
 of Gerald M. Rubin (University of California, Berkeley)  
 using ZAP-cDNA synthesis kit (Stratagene) and Superscript  
 II RT (Life Technologies). Note: this is a NIH\_MGC  
 Library."  
 BASE COUNT 225 a 233 c 252 g 221 t 1 others  
 ORIGIN  
 Query Match 76.9%; Score 318.4; DB 14; Length 932;  
 Best Local Similarity 88.6%; Pred. No. 1.8e-87;  
 Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

QY 3 AAGCTACAGAACAGCTTGCAAGTTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62  
 |||||  
 Db 298 AACTTACAGAACAGCTTGCAAAATTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 356  
 |||||  
 QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCCTTAACACGCTGGACCA 122  
 |||||  
 Db 357 TTGATATCTGGAACATGATGGAAGCTTCCGAGACAATGGCCTTATACACTGGACCA 416  
 |||||  
 QY 123 GCACGAGATCAGGGGTGTCGCCCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182  
 |||||  
 Db 417 CCACCGAGATCA-GTGTGTCGCCCTCGAAACTGTCATCTCTCCATCTACTATCAGTTG 475  
 |||||  
 QY 183 ACAACGGCCTTCCTCTCTACTCACCAGATCAGGGTGGAGACATCCATCAGTCTCTCTCTC 242  
 |||||

Db 476 AACAGCGCCTTCCTCTACTACCAAATAGTGTGGAACAATCTATCAGCCTCCTCCTC 535  
 |||||  
 QY 243 AATTTCATGTCGCCCTGACACAGTGTAGCGCCGAGGCANGTTGACCGTGTCTTCAGNT 302  
 |||||  
 Db 536 AACTTTATGATGTCATATGACAGTGTAGCGCCGAGGCAGTGTGACGGTATTTTCAGTT 595  
 |||||  
 QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362  
 |||||  
 Db 596 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 655  
 |||||  
 QY 363 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTNGGAAAGCTTGACCAAGT 414  
 |||||  
 Db 656 TCCAGATGTCAGATCCCAATGGCTTAATGATGTTTAGCAAGTTTGACCAAGT 707  
 |||||

RESULT 9  
 AL524324 967 bp mRNA linear EST 13-FEB-2001  
 LOCUS AL524324 LTL\_NFL003\_NBC3 Homo sapiens cDNA clone CS0DC006YE12 5  
 DEFINITION prime, mRNA sequence.  
 ACCESSION AL524324  
 VERSION AL524324.1 GI:12787817  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 Li, W.B., Gruber, C., Jesses, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr).

FEATURES  
 Location/Qualifiers  
 1..967  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0DC006YE12"  
 /clone\_lib="LTL\_NFL003\_NBC3"  
 /sex="male"  
 /tissue\_type="neuroblastoma cells"  
 /lab\_host="DH10B"  
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-stranded cDNA was digested with Not I and  
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6  
 vector. Library was normalized. Library was constructed  
 by Life Technologies. Contact : Feng Liang Life  
 Technologies, a division of Invitrogen 9800 Medical Center  
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610  
 8371 Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :  
<http://fulllength.invitrogen.com>"  
 BASE COUNT 229 a 250 c 256 g 227 t 5 others  
 ORIGIN  
 Query Match 76.9%; Score 318.4; DB 9; Length 967;  
 Best Local Similarity 88.6%; Pred. No. 1.8e-87;  
 Matches 365; Conservative 0; Mismatches 45; Indels 2; Gaps 2;

QY 3 AAGCTACAGAACAGCTTGCAAGTTTACCGATTTGTACAGAGCGATGCAACCTTCATCTTG 62  
 |||||  
 Db 322 AACTTACAGAACAGCTTGCAAAATTA-CGATTTGTACAAAACGATGCAACCTTCATCTTG 380  
 |||||  
 QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCCTTAACACGCTGGACCA 122  
 |||||  
 Db 381 TTGATATCTGGAACATGATGGAAGCTTCCGAGACAATGGCCTTATACACTGGACCA 440  
 |||||  
 QY 123 GCACGAGATCAGGGGTGTCGCCCTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182  
 |||||  
 Db 441 CCACCGAGATCA-GTGTGTCGCCCTCGAAACTGTCATCTCTCCATCTACTATCAGTTG 499  
 |||||

QY	193	ACAAGCGCCTTCCTTCTACTACACAGATACAGCGTGGAGCAGTCCATCAGTCTCCTACTC	242
Db	500	AAAACAGCGCCTTCCTTCTACTACACAAATAGTGTGGAAACAATCTATCAGCCTCCTCCTC	559
QY	243	AATTTATGCTGCGCGCTACGACAGTGGAGCGCGAGCGANGTTGACCGTGTTCAGNT	302
Db	560	AACTTATGATTGCTGCATATGACAGTGGAGCGCGAGCGANGTTGACCGTGTTCAGNT	619
QY	303	AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGCAAAATTCAGATACATTTTC	362
Db	620	AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGCAAAATTCAGATACATTTTC	679
QY	363	TCCAGATGTCAGATTCGAATGGCTTAATGATGTTNGGAAAGCTTGACCACT	414
Db	680	TCCAGATGTCAGATTCGAATGGCTTAATGATGTTNGGAAAGCTTGACCACT	731
RF	01LT 10		
	: 818		
DEFINITION	AL555818	896 bp	mrna linear EST 16-FEB-2001
ACCESSION	AL555818	prime, mRNA sequence.	
VERSION	AL555818		
KEYWORDS	AL555818.1	GI:12897911	
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Li, W.B., Gruber, C., Jessee, J., and Polayes, D.		
TITLE	Full-length cDNA libraries and normalization		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Genoscope		
FEATURES	Genoscope - Centre National de Sequencage		
source	BP 191 91006 EVRY cedex - France		
	Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		
	Location/Qualifiers		
	1. .896		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0DK002YG20"		
	/clone_lib="LTI.NFL006.PL2"		
	/tissue_type="placenta"		
	/note="Vector: pcwvSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"		
BASE COUNT	224 a	221 c	227 g
ORIGIN			4 others
Query Match	76.88;	Score 318;	DB 9; Length 896;
Best Local Similarity	88.38;	Pred. No. 2.3e-87;	
Matches 364;	Conservative 1;	Mismatches 45;	Indels 2; Gaps 2;
QY	3	AACGTACAGAACAGCTTGCAAGTTACCGATTGTTACAGAGACGATGCAACCTTCATCTTG	62
Db	261	AACGTACAGAACAGCTTGCAAAATTA-CGATTGTGACAAAACGATGCAACCTTCATCTTG	319
QY	63	TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGGCCCTTAACACGCTGGACCACA	122
Db	320	TTGATATCTGGAACATGATTGAMGCCCTTCGAGACATGGCCCTTAATACACTGGACCATA	379
QY	123	GCACGGAGATCAGCGGTGTCGCCCTGGAGACCGTATCTGCTCCATCTACTATCAGTTG	182
Db	380	CCACCGAGATCA-GTGTGTCGCCCTGGAACACTGCTATCTCCCTCACTACTATCAGTTG	438

```

further details.
e mouse tissues.
FEATURES
    source
        1. .647
            Location/Qualifiers
                /organism="Mus musculus"
                /db_xref="taxon:10090"
                /clone="D030074022"
                /clone_lib="RIKEN full-length enriched, 9 days embryo"
                /dev_stage="9 days embryo"
                /lab_host="DH10B"
                /note="Site.1: Salt; Site.2: BamHI: cDNA library was
                    prepared and sequenced in Mouse Genome Encyclopedia
                    Project of Genome Exploration Research Group in Riken
                    Genomic Sciences Center and Genome Science Laboratory in
                    RIKEN. Division of Experimental Animal Research in Riken
                    contributed to prepare mouse tissues. 1st strand cDNA was
                    primed with a primer 15'

```

prepared by using renatase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATCTCGATTATTAATATCCCCCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified

BASE COUNT	136 a	179 c	204 g	128 t
ORIGIN				
Query Match		75 2%	Score 311 4:	DB 10: Length 647:

Matches 334; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Db	310	AACGTACAGAACAGCTTGCAAGTTA - CGATTGTATACAGAAGCGATGCAACCTTCATCTTG	368
QY	63	TTGATATCTCGAATATGATCGAAGCTTTCGAGACAAATGGCCTTAACACGCTGGACCACA	122
Db	369	TTGATATCTGGAATATGATCGAAGCTTTCGAGACAAATGGCCTTAACACGCTGGACCACA	428

Db 429 GCACGGAGATCA -GGTGTCCCGCTGGAGACGTCATCTCGTCCATCTACTATCATGTT 487  
QY 183 AACAAAGCGCCTTCTTCTTACTACCAAGATCAGCGTGAGCAGTGCCATCAGTCTCCTACTC 242  
Db 488 AACAAAGCGCCTTCTTCTTACTACCAAGATCAGCGTGAGCAGTGCCATCAGTCTCCTACTC 547

QY	243	AAATTTCATGTCGCGCCCTACGACAGTCAGCGCGGACGGCANGTGTGACCGTGTTCAGNT	302
Db	548	AAATTTCATGTCGCGCCCTACGACAGTCAGCGCGGACGGCANGTGTGACCGTGTTCAGTT	607
QY	303	AAAGCTATGTTAGCAACCATGTGTGTTGGAAAAATGCTGG	342
Db	608	AAAGCTATGTTAGCAACCATGTGTGTTGGAAAAATGCTGG	647

RESULT 13  
RG078658

LOCUS	BC078688	590 bp	mRNA	linear	EST 26-JAN-2001
DEFINITION	H3030G07-5 NIA Mouse 15K cDNA Clone Set				
	H3030G07 5', mRNA sequence.				
ACCESSION	BC078688				
VERSION	BC078688				
KEYWORDS	BC078688.1	GI:12561256			
SOURCE	EST.				
ORGANISM	house mouse.				
	Mus musculus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 (bases 1 to 590)				
AUTHORS	Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Jaradat,S.A., Tanaka				
	,T.S., Carter,M.G. and Ko,M.S.H.				
TITLE	Verification and initial annotation of NIA mouse 15K cDNA clone set				

JOURNAL  
COMMENT

Unpublished (2001)  
Other\_ESTs: H3030G07-3  
Contact: George J. Kargul  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdnaelgusun.grc.nia.nih.gov  
This clone set has been freely distributed to the community. Please  
visit <http://lgusun.grc.nia.nih.gov/cdna/15k.html> for details.  
Plate: H3030 row: G column: 07  
Seq primer: -21M13 Reverse  
High quality sequence stop: 590  
POLYA-No.

FEATURES  
source

Location/Qualifiers  
1..590  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="niaEST:H3030G07-5"  
/db\_xref="taxon:10090"  
/clone="H3030G07"  
/clone.lib="NIA Mouse 15K cDNA Clone Set"  
/sex="Clones arrayed from a variety of cDNA libraries"  
/dev\_stage="Clones arrayed from a variety of cDNA libraries"  
/lab\_host="DH10B"  
/note="Vector: pSPORT1; Site\_1: SalI; Site\_2: NotI; This clone is among a rearranged set of 15,247 clones from 11 embryonic cDNA libraries (including preimplantation stage embryos from unfertilized egg to blastocyst, embryonic part of E7.5 embryos, extraembryonic part of E7.5 embryos and E12.5 female mesonephros/gonad) and one newborn ovary cDNA library. Average insert size 1.5 Kb. All source libraries are cloned unidirectionally with Oligo(dT)-Not primers. References include: (1) Genome-wide expression profiling of mid-gestation placenta and embryo using a 15,000 mouse developmental cDNA microarray, 2000, Proc. Natl. Acad. Sci. U S A, 97: 9127-9132; (2) Large-scale cDNA analysis reveals phased gene expression patterns during preimplantation mouse development, 2000, Development, 127: 1737-1749; (3) Genome-wide mapping of unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978."  
141 a 158 c 138 g 153 t

BASE COUNT  
ORIGIN

Query Match 74.2%; Score 307; DB 12; Length 590;  
Best Local Similarity 98.5%; Pred. No. 4.6e-84;  
Matches 318; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 92 CGAGACATGGCTTAACACGCTGGACACAGCAGGAGATCAGGCGTTCGCCGTGGA 151  
DB 1 CGAGACATGGCTTAACACGCTGGACACAGCAGGAGATCA-GCGTGTCCGCTGGA 59  
QY 152 GACCGTCATCTCGTCATCTACTATCAGTTGAACAGCGCCTTCCTTACTCACCAGAT 211  
DB 60 GACCGTCATCTCGTCATCTACTATCAGTTGAACAGCGCCTTCCTTACTCACCAGAT 119  
QY 212 CAGCGTGGAGCAGTCCATCAGTCTCTACTCAATTTTCATGTCGCGCTACGACAGTGA 271  
DB 120 CAGCGTGGAGCAGTCCATCAGTCTCTACTCAATTTTCATGTCGCGCTACGACAGTGA 179  
QY 272 GGGCCGAGCAGTGGACCGTGTTCAGTAAAGCTATGTTAGCAACCATGTGTGGTGG 331  
DB 180 GGGCCGAGCAGTGGACCGTGTTCAGTAAAGCTATGTTAGCAACCATGTGTGGTGG 239  
QY 332 AAAATGCTGGACAATTCAGATACATTTTCCACAGATGTCAGATTCATGCGCTTAAT 391  
DB 240 AAAATGCTGGACAATTCAGATACATTTTCCACAGATGTCAGATTCATGCGCTTAAT 299  
QY 392 GATGTTNGGAAAGCTTGACCACT 414  
||||| |||||||

Db 300 GATGTTTGGAAAGCTTGACCACT 322

RESULT 14  
BG805399  
LOCUS

## DEFINITION

0692-19 Mouse E14.5 retina lambda ZAP II Library Mus musculus cDNA,  
mRNA sequence.

## ACCESSION

600 bp mRNA linear EST 20-DEC-2001  
BG805399

## VERSION

1 GI:17952332

## KEYWORDS

EST.

## SOURCE

house mouse.

## ORGANISM

Mus musculus

## REFERENCE

1 (bases 1 to 600)

## AUTHORS

Mu, X., Zhao, S., Pershad, R., Hsieh, T.-F., Scarpa, A., Wang, S.W.,

## TITLE

White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. and Klein, W.H.

## JOURNAL

Gene expression in the developing mouse retina by EST sequencing

## MEDLINE

and microarray analysis

## COMMENT

Nucleic Acids Res. 29 (24), 4983-4993 (2001)

## CONTACT

21671825

## DEPARTMENT

Contact: Klein WH

## UNIVERSITY

Department of Biochemistry and Molecular Biology

## BOX

University of Texas M.D. Anderson Cancer Center

## TEL

Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA

## FAX

Tel: 713 792 3646

## FEATURES

Fax: 713 790 0329

## source

Location/Qualifiers

1..600

/organism="Mus musculus"

/clone.lib="Mouse E14.5 retina lambda ZAP II Library"

/tissue\_type="neural retina"

/dev\_stage="embryonic day 14.5 post-fertilization"

/note="Vector: pAMP10 (Gibco); Cloned unidirectionally.

Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps

(Mannatis); Cloning Technique: CUA Cloning (Cloneamp,

Life Technologies); Average insert size: 1.8 Kb;

Insertion site: TACGTCCACTGAATCTGAGTG---. Other

information regarding entire library may be found at

[http://pga.swmed.edu/Data/libraries/microarray\\_cdna\\_library.htm](http://pga.swmed.edu/Data/libraries/microarray_cdna_library.htm).

BASE COUNT 130 a 167 c 180 g 123 t

Query Match 73.2%; Score 303; DB 12; Length 600;

Best Local Similarity 98.5%; Pred. No. 8.1e-83;

Matches 325; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY 3 AACGTACAGACAGCTTCCAAAGTTACCGATTGTACAGAGCGGATGCAACCTTCATCTTG 62

DB 273 AACGTACAGACAGCTTCCAAAGTTA-CGATTTGTACAGAGCGGATGCAACCTTCATCTTG 331

QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCTTAACACGCTGGACACACA 122

DB 332 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAATGGCTTAACACGCTGGACACACA 391

QY 123 GCACGGAGATCAGCGGTGTCGCCCTGGAGACCGTCTCTCGTCCATCTACTATCAGTTG 182

DB 392 GCACGGAGATCA-GCGTGTCCCGCTGGAGACCGCTCTCTCGTCCATCTACTATCAGTTG 450

QY 183 AACAGCGCCTTCTCTTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 242

DB 451 AACAGCGCCTTCTCTTACTCACCAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 510

QY 243 AATTTCATGTGCGCGCTACGACAGTGAAGCGCGAGCGANGTTGACCGTGTTCAGTT 302

DB 511 AATTTCATGTGCGCGCTACGACAGTGAAGCGCGAGCGANGTTGACCGTGTTCAGTT 570

QY 303 AAAGCTATGTTAGCAACCATGTGTGGGA 332

DB 571 AAAGCTATGTTAGCAACCATGTGTGGGA 600

RESULT 15  
BG469017  
LOCUS  
DEFINITION BG469017 830 bp mRNA linear EST 21-MAR-2001  
602510757p1 NIH\_MGC\_15 Homo sapiens cDNA clone IMAGE:4644783 5',  
mRNA sequence.  
ACCESSION BG469017  
VERSION BG469017.1 GI:13401202  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 830)  
AUTHORS NIH-MGC <http://mgc.nhl.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-re@mail.nih.gov](mailto:cgapbs-re@mail.nih.gov)  
Tissue Procurement: ATCC  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: NIH Intramural Sequencing Center  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LCM1417 row: b column: 16  
High quality sequence stop: 646.

FEATURES  
source  
1..830  
Location/Qualifiers  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4644783"  
/clone\_lib="NIH\_MGC\_15"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/note="Organ: colon; Vector: pORF7; Site\_1: XhoI; Site\_2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5',  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)"  
BASE COUNT 213 a 207 c 221 g 189 t

Query Match 72.3%; Score 299.4; DB 12; Length 830;  
1st Local Similarity 87.2%; Pred. No. 1.3e-81;  
atches 360; Conservative 0; Mismatches 50; Indels 3; Gaps 3;  
QY 3 AACGTACAGACAGCTTGCAAGTTACCGATTTGTACAGAGGAGGATGCAACCTTCATCTTG 62  
DB 265 AACTTACAGACAGCTGCAATTA-CGATTGTACAAAACAGATGCAACCTTCATCTTG 323  
QY 63 TTGATATCTGGAATATGATCGAAGCTTTCCGAGACAAATGCCCTTAACACGCTGGACCACA 122  
DB 324 TTGATATCTGGAACATGATTGAAGCCCTCCGAGACAAATGCCCTTAACACGCTGGACCATA 383  
QY 123 GCACGGAGATCAGCGGTGTNCCGCGTGGAGACCGTCATCTCGTCCATCTACTATCAGTTG 182  
DB 384 CCACCGAGATCA-GTGTGTCGCGCTCGAAACTGTCATCTCCTCCATCTACTATCAGTTG 442  
QY 183 AACAGCGCTTCCTTCTACTACACAGATCAGCGTGGAGCAGTCCATCAGTCTCCTACTC 242  
DB 443 AACAGCGCTTCCTTCTACTACCAAAATAGTGTGGACAACTCTATCAGCCTCCTCCTC 502  
QY 243 AATTTTCATGTCGCGCCCTTACGACAGTGGAGGCCGAGGCAAGTTGACCGTGTTCAGNT 302  
DB 503 AACTTTATGATTCGTCATATGACATGAGGGCCGAGGCAAGTTGACGGTATTTTCAGTT 562  
QY 303 AAAGCTATGTTAGCAACCATGTGTGGTGGAAAAATGCTGGACAAATTGAGATACATTTTC 362

Db 563 AAAGCTATGTTAGCAACCATGTGTGGTGGAAACAATGCTGGACAATCGAGATATGTTATC 622  
Qy 363 TCCAGATGTCAGATTCGAATGG-CTTAATGATGTTNGGAAAGCTTGACCACT 414  
Db 623 TCCAGATGTCAGATTCGAATGGCCTTCATGATATTCAGCAAAAGTTGACCACT 675  
Search completed: December 1, 2002, 17:42:10  
Job time : 1989 secs